# An Iterated Azéma-Yor Type Embedding for Finitely Many Marginals

Jan Obłój* and Peter Spoida†
Mathematical Institute *and*
Oxford-Man Institute of Quantitative Finance,
University of Oxford

April 2, 2013

**Abstract**

We propose an iterated Azéma-Yor type embedding in the spirit of [Azéma and Yor, 1979] and [Brown et al., 2001a] for any given finite number $n \in \mathbb{N}$ of probability measures $\mu_1, \dots, \mu_n$ which are in convex order and satisfy an additional technical assumption. In particular, our construction reproduces the stopping boundaries obtained in [Madan and Yor, 2002] and [Brown et al., 2001a]. We demonstrate with a counterexample that our technical assumption is necessary and propose an extended embedding for $n = 3$. As a by-product of our analysis we compute the law of the maximum at all the stopping times. This is used in our parallel work [Henry-Labordère et al., 2013] to establish extremal properties of our embedding and develop robust pricing and hedging of Lookback options given prices of call options with multiple intermediate maturities.

## 1 Introduction

The Skorokhod embedding problem (SEP) consists in finding a stopping time $\tau$ such that a given stochastic process $(X_t)$ stopped at $\tau$ has a given distribution $\mu$. For the solution to be useful (and non-trivial) one further requires $\tau$ to be *minimal* (cf. [Obłój, 2004, Sec. 8]) which, when $X$ is a continuous local martingale, is equivalent to $(X_{t \wedge \tau} : t \geqslant 0)$ being a uniformly integrable martingale. The problem, dating back to the original work in [Skorokhod, 1965], has been an

*jan.obloj@maths.ox.ac.uk; `http://www.maths.ox.ac.uk/people/profiles/jan.obloj`
†peter.spoida@maths.ox.ac.uk; `http://www.maths.ox.ac.uk/people/profiles/peter.spoida`

active field of research for nearly 50 years. New solutions often either considered new classes of processes $X$ or focused on finding stopping times $\tau$ with additional optimal properties. In particular, the latter is true for works on SEP motivated by mathematical finance.

The classical approach in quantitative finance postulates a model – a stochastic process on a filtered probability space – and uses the model together with no-arbitrage assumption to deduce the fair price today of an option with payoff $O(X)_T$, where $O(X)_T$ is simply an $\mathbb{F}_T$-measurable random variable. The no-arbitrage assumption implies, up to an equivalent change of measure and discounting, that our stochastic process is a martingale and then the price is given as $\mathbb{E}[O(X)_T]$. However in practice many of such options have known prices quoted in the market. It is therefore natural to pose an inverse problem: given a set of options with market prices what is the range of possible prices for $O(X)_T$ among all consistent models? This logic is often referred to as *robust*, or *model-free*, approach. If we assume that market quoted prices for options with payoff $(X_T - K)^+$ (call options) are given for all $K$ then this is equivalent to fixing the marginal law $X_T \sim \mu$. If we further assume that $X$ has continuous paths then consistent models correspond to all continuous martingales with a given marginal at time $T$. These can naturally be represented as time-changes of Brownian motion $B$ and, as long as $O(X)_T$ is invariant under time change, the robust pricing question above boils down to finding the range of $\mathbb{E}[O(B)_\tau]$ among all solutions to the Skorokhod embedding problem.

The above link between SEP and robust pricing and hedging was pioneered by [Hobson, 1998] who considered Lookback options. Barrier options were subsequently dealt with by [Brown et al., 2001b]. More recently, [Cox and Obłój, 2011b, Cox and Obłój, 2011a] considered the case of double touch/no-touch barrier options, [Hobson and Neuberger, 2012] looked at forward starting straddles and analysis for variance options was undertaken in [Cox and Wang, 2013]. We refer to [Hobson, 2010] and [Obłój, 2010] for an exposition of the main ideas and more references.

All the above works assumed that prices of call options with only one maturity are given, or equivalently that one marginal at the final time $T$ is fixed[1]. Meanwhile, in practice, often prices of call options with several intermediate maturities are also available. Mathematically, this translates into optimising $\mathbb{E}[O(B)_{\tau_n}]$ over all solutions $\tau_1 \leqslant \cdots \leqslant \tau_n$ to a multi-marginal Skorokhod embedding problem: $B_{\tau_i} \sim \mu_i$, $1 \leqslant i \leqslant n$, and $(B_{t\wedge\tau_n})_{t\geqslant 0}$ is a uniformly integrable martingale, where $\mu_1, \ldots, \mu_n$ is a given sequence of centred probability measures non-decreasing in the convex order. As shown by [Strassen, 1965], the convex ordering of measures is both necessary and sufficient for a solution to exist.

[1] With the exception of [Hobson and Neuberger, 2012] where a second marginal corresponding to the starting date of the option is also fixed leading to a one-marginal SEP but with a non-trivial starting law.

This problem turns out to be significantly more complex that the special case $n = 1$. In fact, we are not aware of any previous works which would solve it for non-trivial examples of $O$. The only exception is [Brown et al., 2001a] where the authors constructed a generalisation of the Azéma–Yor embedding to two marginals and proved its optimal properties thus solving the problem for $n = 2$ and $O(B)_\tau = \phi(\sup_{t\leqslant\tau} B_t)$ for an increasing function $\phi$. In this paper we extend this further and treat the case of arbitrary $n \in \mathbb{N}$. We prove the embedding property for a general sequence of measures which satisfy an additional technical assumption.

We started our quest for a general $n$-marginal optimal embedding by computing the value function $\sup \mathbb{E}\left[\phi(\sup_{t\leqslant\tau_n} B_t)\right]$ among all solutions to the $n$-marginal SEP. This was achieved using stochastic control methods, developed first for $n = 1$ in [Galichon et al., 2013], and is reported in a companion paper [Henry-Labordère et al., 2013]. Knowing the value function we could start guessing the form of the optimiser and this led to the present paper. Consequently the optimal properties of our embedding, namely that it indeed achieves the value function in question, are shown in [Henry-Labordère et al., 2013]. In fact we give two proofs in that paper, one via stochastic control methods and another one by constructing appropriate pathwise inequalities.

The reason why the $n$-marginal problem is much more involved than the one-marginal one is that while it might be easy to iterate a one-marginal solution, there is no guarantee that such a simple iteration of optimal embeddings would be globally optimal. Indeed, this is usually not the case. Consider the Azéma–Yor embedding [Azéma and Yor, 1979] which consists of a first exit time for the joint process $(B_t, \bar{B}_t)_{t\geqslant 0}$, where $\bar{B}_t = \sup_{s\leqslant t} B_s$. More precisely, their solution $\tau^{\mathrm{AY}} = \inf\left\{t \geqslant 0 : B_t \leqslant \xi_\mu(\bar{B}_t)\right\}$ leads to a functional relation $B_{\tau^{\mathrm{AY}}} = \xi_\mu(\bar{B}_{\tau^{\mathrm{AY}}})$. This then translates into the extremal property that the distribution of $\bar{B}_{\tau^{\mathrm{AY}}}$ is maximized in stochastic order amongst all solutions to SEP for $\mu$, i.e. for all $y$,

$$\mathbb{P}\left[\bar{B}_{\tau^{\mathrm{AY}}} \geqslant y\right] = \sup\left\{\mathbb{P}\left[\bar{B}_\rho \geqslant y\right] : \rho \text{ s.t. } B_\rho \sim \mu, (B_{t\wedge\rho}) \text{ is UI }\right\}.$$

It is not hard to generalise the Azéma-Yor embedding to a non-tirivial starting law, see [Obłój, 2004, Sec. 5]. Consequently we can find $\eta_i$ such that $\tau_i = \inf\left\{t \geqslant \tau_{i-1} : B_t \leqslant \eta_i(\sup_{\tau_{i-1}\leqslant s\leqslant t} B_s)\right\}$ solve the $n$-marginal SEP. However this construction will maximise stochastically the distributions of $\sup_{\tau_{i-1}\leqslant t\leqslant\tau_i} B_t$ but not of the global maximum $\bar{B}_{\tau_n}$.

Instead, we propose an "iteration" of the Azéma-Yor embedding in the following sense. Each $\tau_{i+1}$ is still a first exit for $(B_t, \bar{B}_t)_{t\geqslant\tau_i}$ which is designed in such a way as to obtain a "strong relation" between $B_{\tau_{i+1}}$ and $\bar{B}_{\tau_{i+1}}$, ideally a functional relation. Under our technical assumption about the measures $\mu_1, \ldots, \mu_n$, Assumption A, we describe this relation in detail in Lemma 3.1 which is key for the pathwise proof of the optimal properties of our embedding in Section 4 in [Henry-Labordère et al., 2013].

For $n = 2$ our work recovers the results in [Brown et al., 2001b]. We also recover the trivial case $\tau_i = \tau^{AY}_{\mu_i}$ which happens when $\xi_{\mu_i} \leqslant \xi_{\mu_{i+1}}$, we refer to [Madan and Yor, 2002] who in particular then investigate properties of the arising time-changed process. However, as a counterexample shows, our construction does not work for any laws $\mu_1, \ldots, \mu_n$ which are in convex order. A more involved analysis is required to obtain an explicit embedding when our Assumption A fails. We only sketch the appropriate arguments for the case $n = 3$ as this case does not seem to be of any practical relevance.

The remainder of the article is organized as follows. In Section 2 we explain the main quantities for the embedding and state the main result. We also present the restriction on the measures $\mu_1, \ldots, \mu_n$ which we require for our construction to work (Assumption A). In Section 3 we prove the main result and Section 4 provides a discussion of extensions together with comments on Assumption A. The proof of an important but technical lemma is relegated to the Appendix.

# 2 Main Result

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space satisfying the usual hypothesis and $B$ an $\mathbb{F}$–continuous local martingale, $B_0 = 0$, $\langle B \rangle_\infty = \infty$ a.s. and $B$ has no intervals of constancy a.s. We denote $\bar{B}_t := \sup_{s \leqslant t} B_t$. We are primarily interested in the case when $B$ is a standard Brownian motion and it is convenient to keep this example in mind, hence the notation. We allow for more generality as this introduces no changes to the proof of statements.

## 2.1 Definitions

The following definition will be crucial in the remainder of the article. We define the stopping boundaries $\xi_1, \ldots, \xi_n$ for our iterated Azéma-Yor type embedding together with quantities $K_1, \ldots, K_n$ which will be later linked to the law of the maximum at subsequent stopping times.

**Definition 2.1.** *Fix $n \in \mathbb{N}$. For convenience we set*

$$c_0 \equiv 0, \quad K_0 \equiv 0, \quad \xi_0(y) = -\infty. \tag{2.1}$$

*For $\zeta \in \mathbb{R}$ and $i = 1, \ldots, n$ we write*

$$c_i(\zeta) := \int_{\mathbb{R}} (x - \zeta)^+ \, \mu_i(\mathrm{d}x). \tag{2.2}$$

*Let $y \geqslant 0$ and assume that for $i = 1, \ldots, n-1$ the quantities $\xi_i, K_i, \imath_i$ and $\jmath_i$ are already defined. Then we define*

$$\begin{aligned} &\imath_n(\cdot; y) : (-\infty, y] \to \{0, 1, \ldots, n-1\}, \\ &\imath_n(\zeta; y) := \max\{k \in \{0, 1, \ldots, n-1\} : \xi_k(y) < \zeta\}, \end{aligned} \tag{2.3}$$

*and*

$$\xi_n(y) := \sup\left\{\arg\inf_{\zeta<y}\left\{\frac{c_n(\zeta)}{y-\zeta} - \left[\frac{c_{\imath_n(\zeta;y)}(\zeta)}{y-\zeta} - K_{\imath_n(\zeta;y)}(y)\right]\right\}\right\}. \tag{2.4}$$

*With*

$$\jmath_n(y) := \imath_n(\xi_n(y);y) \tag{2.5}$$

*we set*

$$K_n(y) := \frac{1}{y-\xi_n(y)}\left\{c_n(\xi_n(y)) - \left[c_{\jmath_n(y)}(\xi_n(y)) - (y-\xi_n(y))K_{\jmath_n(y)}(y)\right]\right\}. \tag{2.6}$$

**Definition 2.2** (Embedding)**.** *Set $\tau \equiv 0$ and for $i = 1, \ldots, n$ define*

$$\tau_i := \begin{cases} \inf\left\{t \geqslant \tau_{i-1} : B_t \leqslant \xi_i(\bar{B}_t)\right\} & \text{if } B_{\tau_{i-1}} > \xi_i(\bar{B}_{\tau_{i-1}}), \quad (2.7a) \\ \tau_{i-1} & \text{else.} \quad (2.7b) \end{cases}$$

*Remark* 2.3 (Properties of $\imath_n$). Recalling the definition of $\imath_n$, cf. (2.3), we observe for later use that for $y \geqslant 0$

$$\imath_n(\cdot;y) \text{ is left-continuous and has at most } n-1 \text{ jumps} \tag{2.8}$$

and for $x \in \mathbb{R}$

$$\imath_n(x;\cdot) \text{ is right-continuous and has at most } n-1 \text{ jumps.} \tag{2.9}$$

Figure 2.1 illustrates a set of possible stopping boundaries in the case of $n = 3$. If Assumption A is in place, see Section 2.2, we will show that the stopping boundaries are continuous (except possibly for $i = 1$) and non-decreasing, cf. Section 2.5.

The $n^{\text{th}}$ stopping boundary $\xi_n$ is obtained from an optimization problem which features $\xi_1, \ldots, \xi_{n-1}$ and $K_1, \ldots, K_{n-1}$. $K_n(y)$ is the value of the objective function at the optimal value $\xi_n(y)$. Note that all previously defined stopping boundaries $\xi_1, \ldots, \xi_{n-1}$ and the quantities $K_1, \ldots, K_{n-1}$ remain unchanged.

Denote the right endpoint of the support of the measure $\mu_i$ by

$$r_{\mu_i} := \inf\{x : \mu_i((x,\infty)) = 0\} \tag{2.10}$$

and the barycentre function of $\mu_i$ by

$$b_i(x) := \frac{\int_{[x,\infty)} u \mathrm{d}\mu_i(u)}{\bar{\mu}_i([x,\infty))}\mathbb{1}_{\{x<r_{\mu_i}\}} + x\mathbb{1}_{\{x\geqslant r_{\mu_i}\}}, \tag{2.11}$$

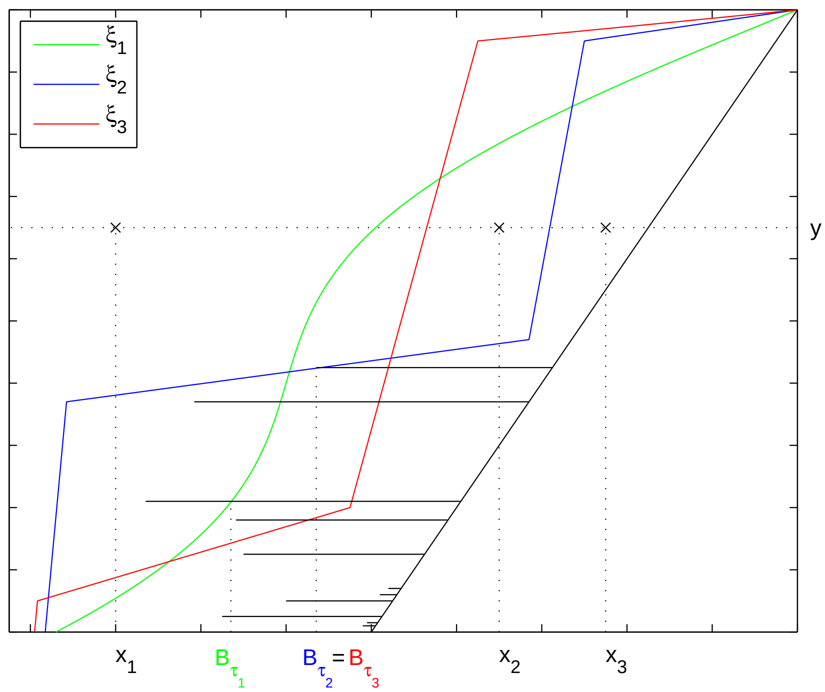


Figure 2.1: We illustrate possible stopping boundaries $\xi_1, \xi_2, \xi_3$. The horizontal lines represent a sample path of the process $\big(B_t, \bar{B}_t\big)$ where the $x$-axis is the value of $B$ and the $y$-axis the value of $\bar{B}$. Each horizontal segment is an excursion of $B$ away from its maximum $\bar{B}$. The first stopping time $\tau_1$ is found when the process first hits $\xi_1$. Since $\xi_1(\bar{B}_{\tau_1}) > \xi_2(\bar{B}_{\tau_1})$ the process continues and targets $\xi_2$. The stopping time $\tau_2$ is found when the process first hits $\xi_2$. Since $\xi_2(\bar{B}_{\tau_2}) \leqslant \xi_3(\bar{B}_{\tau_2})$ we get $\tau_3 = \tau_2$. For the $y$ we fixed we have $\imath_3(x_1; y) = 0, \imath_3(x_2, y) = 1, \imath_3(x_3; y) = 2$.

As shown by [Brown et al., 2001b], the right-continuous inverse of $b_i$, denoted by $b_i^{-1}$, can be represented as

$$b_i^{-1}(y) = \sup\left\{\arg\inf_{\zeta<y} \frac{c_i(\zeta)}{y-\zeta}\right\}. \tag{2.12}$$

It is clear and has been studied in more detail by [Madan and Yor, 2002] that if the sequence of barycentre functions is increasing in $i$, then the intermediate law constraints do not have an impact on the corresponding iterated Azéma-Yor embedding. However, in general the barycentre functions will not be increasing in $i$, cf. [Brown et al., 2001a], and hence will affect the embedding. We think of $\jmath_n(y)$ as the index of the last law $\mu_i, i < n$, which represents, locally at level of maximum $y$, a *binding constraint* for the embedding. As compared to the optimization from which $b_n^{-1}$ is obtained, cf. (2.12), the optimization from which $\xi_n$ is obtained, cf. (2.4), has a penalty term.

## 2.2 Restrictions on Measures

Throughout the article we will denote the left- and right-limit of a function $f$ at $x$ (if it exists) by $f(x-)$ and $f(x+)$, respectively.

For $n \in \mathbb{N}$ and $y \geqslant 0$ define inductively the mappings

$$\begin{aligned} &c^n(\cdot, y) : (-\infty, y] \to \mathbb{R} \\ &x \mapsto c^n(x, y) := c_n(x) - \left[c_{\imath_n(x;y)}(x) - (y-x)K_{\imath_n(x;y)}(y)\right]. \end{aligned} \tag{2.13}$$

It follows that the minimization problem in (2.4) is equivalent to the following minimization problem,

$$\xi_n(y) \in \arg\min_{\zeta \leqslant y} \frac{c^n(\zeta, y)}{y - \zeta}, \tag{2.14}$$

where we observe that

$$\begin{aligned} \left.\frac{c^n(\zeta, y)}{y-\zeta}\right|_{\zeta = y} &:= \lim_{\zeta \uparrow y} \frac{c^n(\zeta, y)}{y - \zeta} \\ &= \begin{cases} -c_n'(y-) + c_{\imath_n(y;y)}'(y-) + K_{\imath_n(y;y)}(y) & \text{if } c_n(y) = c_{\imath_n(y,y)}(y), \\ +\infty & \text{else,} \end{cases} \end{aligned} \tag{2.15}$$

and note that existence in (2.14) can be deduced from the – a priori – piecewise continuity of $c^n(\cdot, y)$, the fact that $c^n \geqslant 0$ and the property that $\zeta \mapsto \frac{c^n(\zeta,y)}{y-\zeta} = \frac{c_n(\zeta)}{y-\zeta}$ for $\zeta$ sufficiently small, which is a non-increasing function.

For later use observe

$$\min_{i \leqslant n} b_i^{-1}(y) \leqslant \xi_n(y) \leqslant y \tag{2.16}$$

which follows from the definition of $\xi_n$, cf. (2.4), and where $b_i^{-1}$ denotes the right-continuous inverse of the barycentre function $b_i$, cf. (2.12).

**Assumption A** (Restriction on Measures)**.** *Recall definitions in* (2.1)–(2.2), (2.13) *and* (2.10). *We impose the following restrictions on* $\mu_1, \ldots, \mu_n$*:*

*(i)* $\int |x| \mu_i(\mathrm{d}x) < \infty$ *with* $\int x \mu_i(\mathrm{d}x) = 0$ *and* $c_{i-1} \leqslant c_i$ *for all* $1 \leqslant i \leqslant n$*,*

*(ii) for all* $2 \leqslant i \leqslant n$ *and all* $0 \leqslant y \leqslant r_{\mu_i}$ *the mapping*

$$(-\infty, y] \to \mathbb{R} \cup \{+\infty\}, \quad \zeta \mapsto \frac{c^i(\zeta, y)}{y - \zeta} \quad \textit{has a unique minimizer } \zeta^\star \tag{2.17}$$

*and*

$$c_i(y) > c_{\imath_i(y;y)}(y) \qquad \textit{whenever } \zeta^\star < y. \tag{2.18}$$

*Remark* 2.4 (Assumption A). The condition that the call prices are non-decreasing in maturity

$$c_i \leqslant c_{i+1}, \qquad i = 1, \ldots, n-1, \tag{2.19}$$

can be rephrased by saying that $\mu_1, \ldots, \mu_n$ are non-decreasing in the convex order. Condition $(i)$ in the Assumption is the necessary and sufficient condition for a uniformly integrable martingale with these marginals to exist, as shown by e.g. [Strassen, 1965].

Note that if (2.19) holds with strict inequality then (2.18) is automatically satisfied.

*Remark* 2.5 (Discontinuity of $\xi_1$). Note that Assumption A(ii) does not require that the mapping

$$\zeta \mapsto \frac{c^1(\zeta, y)}{y - \zeta} = \frac{c_1(\zeta, y)}{y - \zeta} \tag{2.20}$$

has a unique minimizer. It may happen that there is an interval of minimizers and then $\xi_1$ is discontinuous at such $y$.

## 2.3 Main Result

Our main result shows how to iteratively define an embedding of $(\mu_1, \ldots, \mu_n)$ in the spirit of [Azéma and Yor, 1979] and [Brown et al., 2001a] if Assumption A is in place.

**Theorem 2.6** (Main Result). *Let $n \in \mathbb{N}$ and $\mu_1, \ldots, \mu_n$ be probability measures on $\mathbb{R}$ which satisfy Assumption A in Section 2.2. Recall Definitions 2.1 and 2.2. Then $\tau_i < \infty$, $B_{\tau_i} \sim \mu_i$ for all $i = 1, \ldots, n$ and $(B_{\tau_n \wedge t})_{t \geqslant 0}$ is a uniformly integrable martingale.*

*In addition, we have for $y \geqslant 0$ and $i = 1, \ldots, n$,*

$$\mathbb{P}\left[\bar{B}_{\tau_i} \geqslant y\right] = K_i(y) \tag{2.21}$$

*where $K_i$ is defined in* (2.6).

*Remark* 2.7 (Inductive Nature). It is important to observe that $\xi_i$ and therefore also $\tau_i$, only depend on $\mu_1, \ldots, \mu_i$. This gives an iterative structure allowing to "add one marginal at a time" and enables us to naturally prove the Theorem by induction on $n$.

*Remark* 2.8 (Minimality). Since all $\tau_i$ are such that $(B_{t \wedge \tau_i})_{t \geqslant 0}$ is a uniformly integrable martingale it follows from [Monroe, 1972] that all $\tau_i$ are *minimal*.

## 2.4 Examples

Examples 2.9 and 2.10, respectively, show that we recover the stopping boundaries obtained in [Madan and Yor, 2002] and [Brown et al., 2001a], respectively. In particular the case $n = 1$ corresponds to the Azéma-Yor solution [Azéma and Yor, 1979].

**Example 2.9** ( [Madan and Yor, 2002]). *Recall the definition of the barycentre function $b_i$ from* (2.11). *[Madan and Yor, 2002] consider the "increasing mean residial value" case, i.e.*

$$b_1 \leqslant b_2 \leqslant \cdots \leqslant b_n. \tag{2.22}$$

*We will now show that our main result reproduces their result if Assumption A is in place. In fact, as can be seen below, our definitions of $\xi_i$ and $K_i$, cf.* (2.4) *and* (2.6), *respectively, reproduce the correct stopping boundaries in the general case, showing that Assumption A is not necessary, cf. also Section 4. More precisely, we have*

$$\xi_i = b_i^{-1}, \quad K_i(y) = \frac{c_i(b_i^{-1}(y))}{y - b_i^{-1}(y)} =: \mu_i^{\mathrm{HL}}([y, \infty)), \qquad i = 1, \ldots, n, \tag{2.23}$$

*where $b_i^{-1}$ denotes the right-continuous inverse of $b_i$ and $\mu_i^{\mathrm{HL}}$ is the Hardy-Littlewood transform of $\mu_i$, cf. [Carraro et al., 2012].*

*Clearly, the claim is true for $i = 1$. Let us assume that the claim holds for all $i \leqslant n - 1$. Now, the optimization problem for $\xi_n$ in* (2.4) *becomes*

$$\begin{aligned}
\xi_n(y) &\in \arg\min_{\zeta \leqslant y} \left\{ \frac{c_n(\zeta)}{y - \zeta} - \mathbb{1}_{\{\zeta > b_{n-1}^{-1}(y)\}} \left[ \frac{c_{n-1}(\zeta)}{y - \zeta} - \frac{c_{n-1}(b_{n-1}^{-1}(y))}{y - b_{n-1}^{-1}(y)} \right] \right\} \\
&\in \arg\min_{\zeta \leqslant y} \left\{ \min_{\zeta \leqslant b_{n-1}^{-1}(y)} \frac{c_n(\zeta)}{y - \zeta}, \min_{\zeta \geqslant b_{n-1}^{-1}(y)} \left( \frac{c_n(\zeta)}{y - \zeta} - \left[ \frac{c_{n-1}(\zeta)}{y - \zeta} - \frac{c_{n-1}(b_{n-1}^{-1}(y))}{y - b_{n-1}^{-1}(y)} \right] \right) \right\}.
\end{aligned}$$

*It is clear that the first minimum is $A_1 = \frac{c_n(b_n^{-1}(y))}{y - b_n^{-1}(y)}$ since $b_n^{-1}(y) \leqslant b_{n-1}^{-1}(y)$.*

*As for the second minimum, we set*

$$F(\zeta) := \frac{c_n(\zeta)}{y - \zeta} - \left( \frac{c_{n-1}(\zeta)}{y - \zeta} - \frac{c_n\left(b_{n-1}^{-1}(y)\right)}{y - b_{n-1}^{-1}(y)} \right)$$

*and we see by direct calculation that*

$$\begin{aligned}
(y - \zeta)^2 F'(\zeta) &= (b_n(\zeta) - y)\mu_n\left([\zeta, \infty)\right) - (b_{n-1}(\zeta) - y)\mu_{n-1}\left([\zeta, \infty)\right) \\
&= c_n(\zeta) \frac{b_n(\zeta) - y}{b_n(\zeta) - \zeta} - c_{n-1}(\zeta) \frac{b_{n-1}(\zeta) - y}{b_{n-1}(\zeta) - \zeta},
\end{aligned}$$

*where derivatives exist for almost all $\zeta \in \mathbb{R}$. By (2.22), we conclude therefore*

$$(y-\zeta)^2 F'(\zeta) \geqslant (c_n(\zeta) - c_{n-1}(\zeta)) \frac{b_{n-1}(\zeta) - y}{b_{n-1}(\zeta) - \zeta} \geqslant 0,$$

*where the last inequality follows from the non-decrease of the $\mu_i$'s in the convex order. Hence $F$ is non-decreasing, and it follows that it attains its minimum at the left boundary, i.e.* $A_2 = \frac{c_{i+1}(b_i^{-1}(y))}{y-b_i^{-1}(y)} - \left[\frac{c_i(b_i^{-1}(y))}{y-b_i^{-1}(y)} - \frac{c_i(b_i^{-1}(y))}{y-b_i^{-1}(y)}\right] = \frac{c_{i+1}(b_i^{-1})(y)}{y-b_i^{-1}(y)}$. *Consequently,* $\min\{A_1, A_2\} = A_1$ *and* (2.23) *follows.*

**Example 2.10** ( [Brown et al., 2001a])**.** *In the case of $n = 2$ our definition of $\xi_1$ and $\xi_2$ clearly recovers the stopping boundaries in the main result of [Brown et al., 2001a]. However, our embedding is not as general as their embedding because we enforce Assumption A, see also the discussion in Section 4.*

**Example 2.11** (Locally no Constraints)**.** *In general we have*

$$K_n(y) \leqslant \mu_n^{\mathrm{HL}}([y,\infty)). \tag{2.24}$$

*However, if*

$$\xi_n(y) = b_n^{-1}(y) \tag{2.25}$$

*for some $y \geqslant 0$ then it follows from Theorem 2.6 that*

$$K_n(y) = \frac{c_n(b_n^{-1}(y))}{y - b_n^{-1}(y)} = \mu_n^{\mathrm{HL}}([y,\infty)), \tag{2.26}$$

*i.e. locally at level of maximum $y$ the intermediate law constraints have no impact on the distribution of the terminal maximum as compared with the Azéma-Yor embedding.*

## 2.5 Properties of $\xi_n$ and $K_n$

Under Assumption A we establish the continuity of $\xi_n$ for $n \geqslant 2$, cf. Lemma 2.12, and prove monotonicity of $\xi_n$ for $n \geqslant 1$, cf. Lemma 2.13. In Lemma 2.14 we derive an ODE for $K_n$ which will be later used to identify the distribution of the maximum of the embedding from Definition 2.2.

Let $n_1 < n_2$. Recalling Remark 2.7 it follows that the embedding of the first $n_1$ marginals in the $n_2$-marginals embedding problem coincides with the $n_1$-marginals embedding problem. Hence it is natural to prove the Lemma by induction over the number of marginals $n$.

**Lemma 2.12** (Continuity of $\xi_n$)**.** *Let $n \geqslant 2$ and let Assumption A hold. Set*

$$\Delta := \{(x,y) \in \mathbb{R} \times \mathbb{R}_+ : x < y\}. \tag{2.27}$$

*Then the mappings*

$$c^n : \Delta \to \mathbb{R}, \quad (x,y) \mapsto c^n(x,y), \tag{2.28}$$
$$\xi_n : \mathbb{R}_+ \to \mathbb{R}, \quad y \quad \mapsto \xi_n(y) \tag{2.29}$$

*are continuous.*

*Proof.* We prove the claim by induction over $n$. Let us start with the induction basis $n = 1, 2$. Continuity of $c^1$ is the same as continuity of $c_1$ and continuity of $c^2$ is proven in [Brown et al., 2001a], cf. Lemma 3.5 therein. As for continuity of $\xi_2$ we note that our Assumption A(ii) precisely rules out discontinuities of $\xi_2$ as seen in Section 3.5 of [Brown et al., 2001a]. By induction hypothesis we assume continuity of $c^1, \dots, c^{n-1}$ and $\xi_2, \dots, \xi_{n-1}$.

The only possibility that a discontinuity of $c^n$ can occur is when the index $\imath_n$ changes. This only happens at $(x,y) = (\xi_k(y), y)$ for some $k < n$, or, in the case that $y$ is a discontinuity of $\xi_1$, at $(x,y)$ where $x \in [\xi_1(y-), \xi_1(y+)]$. We prove continuity at $(x,y)$.

Consider first the following cases:

$$\begin{array}{llll} & \text{if } x = \xi_k(y) & \text{then} \quad x \neq \xi_j(y) & \text{for all } j \neq k,\ j < n, \quad (2.30) \\ \text{or,} & \text{if } x \in [\xi_1(y-), \xi_1(y+)] & \text{then} \quad x \neq \xi_j(y) & \text{for all } j \neq 1,\ j < n. \quad (2.31) \end{array}$$

Note that in case (2.31) we have from Remark 2.5

$$K_1(y) = \frac{c_1(x)}{y - x} \qquad \text{for all } x \in [\xi_1(y-), \xi_1(y+)]. \tag{2.32}$$

We will call a point $(x,y)$ to be "to the right of $\xi_k$" if $\xi_k(y) < x$ and "to the left of $\xi_k$" if $\xi_k(y) \geqslant x$. From (2.30) and (2.31) it follows that there exists an $\epsilon > 0$ such that each point $(\tilde{x}, \tilde{y})$ in the $\epsilon$-neighbourhood of $(x,y)$ is either to the left or to the right of $\xi_k$ and there are no other boundaries in this $\epsilon$-neighbourhood, in particular

$$k = \imath_n(x_r; y_r), \qquad j = \imath_n(x_l; y_l) = \imath_{\imath_n(x_r;y_r)}(x_r; y_r), \tag{2.33}$$

where $(x_r, y_r)$ is in the $\epsilon$-neighbourhood of $(x,y)$ and to the right of $\xi_k$ and $(x_l, y_l)$ is in the $\epsilon$-neighbourhood of $(x,y)$ and to the left of $\xi_k$.

If $x < y$, we have by induction hypothesis

$$\begin{aligned}
c^n(x_r, y_r) = \quad & c_n(x_r) - \{c_k(x_r) - (y_r - x_r)K_k(y_r)\} && (2.34)\\
\xrightarrow[\text{from the right}]{(x_r,y_r)\to(x,y)} \quad & c_n(x) - \{c_k(x) - (y-x)K_k(y)\} \\
\overset{(2.6)}{\underset{(2.30)-(2.33)}{=}} \quad & c_n(x) - \left\{c_k(x) - \frac{y-x}{y-x}\Big(c_k(x) - [c_j(x) - (y-x)K_j(y)]\Big)\right\} \\
= \quad & c_n(x) - [c_j(x) - (y-x)K_j(y)] \\
\overset{(2.13)}{=} \quad & c^n(x,y) && (2.35)\\
= \quad & c_n(x) - [c_j(x) - (y-x)K_j(y)] \\
\xleftarrow[\text{from the left}]{(x_l,y_l)\to(x,y)} \quad & c_n(x_l) - \{c_j(x_l) - (y_l - x_l)K_j(y_l)\} = c^n(x_l, y_l). && (2.36)
\end{aligned}$$

From (2.34), (2.35) and (2.36) continuity of $c^n$ follows for any sequence $(x_n, y_n) \to (x, y)$. We now extend the above argument to the situation when $x = y$ which establishes left-continuity of $c^n$ at $(y, y)$. In this case we have $x = \xi_k(y) = y$. For this to hold we must have $c_k(y) = c_j(y)$. Using boundedness of $K_i$ for $i < n$ shows that (2.34), (2.35) and (2.36) converge to each other.

To relax (2.30) and (2.31) we successively write out $K_k, K_j, \dots,$ until the assumption of the first case holds true and then, successively, apply the special case.

It remains to prove continuity of $\xi_n$ which we prove by contradiction. Assume there exist $\epsilon > 0$ and $y \geqslant 0$ such that for all $\delta > 0$ there exists a $y' \in (y, y + \delta)$ such that $|\xi_n(y) - \xi_n(y')| > \epsilon$. By (2.16) the limit of $\xi_n(y')$ as $y' \downarrow y$ exists at least along some subsequence and we denote it by $\tilde{\xi}_n$. By assumption $\tilde{\xi}_n \neq \xi_n(y)$.

Consider first the case that $\xi_n(y) < y$ and $\tilde{\xi}_n < y$. Using continuity of $c^n$ we deduce $\frac{c^n(\xi_n(y'),y')}{y'-\xi_n(y')} \to \frac{c^n(\tilde{\xi}_n,y)}{y-\tilde{\xi}_n}$ as $y' \to y$.

Now, if

$$\frac{c^n(\tilde{\xi}_n, y)}{y - \tilde{\xi}_n} \neq \frac{c^n(\xi_n(y), y)}{y - \xi_n(y)} \tag{2.37}$$

then we obtain a contradiction to the optimality of either $\xi_n(y)$ or some $\xi_n(y')$ for $y'$ close enough to $y$ by continuity of $c^n$. If

$$\frac{c^n(\tilde{\xi}_n, y)}{y - \tilde{\xi}_n} = \frac{c^n(\xi_n(y), y)}{y - \xi_n(y)} \tag{2.38}$$

we obtain a contradiction to Assumption A(ii).

We now consider the case that either $\xi_n(y) = y$ or $\tilde{\xi}_n = y$. The case $\xi_n(y) < y$ and $\tilde{\xi}_n = y$ is ruled out by condition (2.18) from Assumption A(ii): Indeed, for the sequence $\left(K_n(y') = \frac{c^n(\xi_n(y'),y')}{y'-\xi_n(y')}\right)$ to be bounded we must have $c^n(\xi_n(y'), y') \to 0$. Recalling the left-continuity of $c^n$ at $(y, y)$ implies $c_n(y) = c_{\imath_n(y;y)}(y)$.

The case $\xi_n(y) = y$ and $\tilde{\xi}_n < y$ follows as above by distinguishing the cases (2.37) and (2.38) and by recalling (2.15) and the left-continuity of $c^n$ at $(y, y)$. $\square$

**Lemma 2.13** (Monotonicity of $\xi_n$). *Let $n \in \mathbb{N}$ and let Assumption A hold. Then the mapping*

$$\xi_n : \mathbb{R}_+ \to \mathbb{R}, \quad y \mapsto \xi_n(y) \tag{2.39}$$

*is non-decreasing.*

*Proof.* The claim for $n = 1, 2$ follows from [Brown et al., 2001a]. Assume by induction hypothesis that we have proven monotonicity of $\xi_1, \ldots, \xi_{n-1}$.

We follow closely the arguments in [Brown et al., 2001a, Lemma 3.2]. Since $\xi_n$ is continuous it is enough to prove monotonicity at almost every $y \geqslant 0$. The set of $y$'s which are a discontinuity of $\xi_1$ is a null-set, and hence we can exclude all such $y$'s. In the following we fix a $y$ where $\xi_1, \ldots, \xi_n$ are continuous.

We will first consider the case when $\xi_n(y) \neq \xi_j(y)$ for all $j < n$. By continuity of $\xi_n$ it follows that there is an $\epsilon > 0$ such that

$$\xi_n(\tilde{y}) \neq \xi_j(\tilde{y}) \text{ and } \ell := \jmath_n(y) = \jmath_n(\tilde{y}) \text{ for all } \tilde{y} \in (y - \epsilon, y + \epsilon) \text{ and } j < n, \tag{2.40}$$

and furthermore

$$(\xi_n(\tilde{y}), \tilde{y}) \in (\xi_n(y) - \epsilon, \xi_n(y) + \epsilon) \times (y - \epsilon, y + \epsilon). \tag{2.41}$$

Let $l_1$ denote a supporting tangent to $c^n(\cdot, y)$ at $\xi_n(y)$ which goes through the $x$-axis at $y$, i.e.

$$l_1(x) = c^n(\xi_n(y), y) + (x - \xi_n(y))(D - K_\ell(y)),$$

where $D$ lies between the left- and right-derivatives of $c_n - c_\ell$ at $\xi_n(y)$. Using that $l_1(y) = 0$ we can write

$$D - K_\ell(y) = -\frac{c^n(\xi_n(y), y)}{y - \xi_n(y)} \overset{(2.13)}{=} -\frac{c_n(\xi_n(y)) - c_\ell(\xi_n(y))}{y - \xi_n(y)} - K_\ell(y)$$

and thus by (2.19)

$$D \leqslant 0. \tag{2.42}$$

We also have

$$l_1(y + \delta) = \delta(D - K_\ell(y)). \tag{2.43}$$

Choose $\delta \in (0, \epsilon)$ sufficiently small. Our goal is to prove $\xi_n(y + \delta) \geqslant \xi_n(y)$. Recall that $\xi_n(y + \delta)$ is determined from $y + \delta$ and $c^n(\cdot, y + \delta)$ only. Since we

know that $\xi_n(y+\delta) \in (\xi_n(y)-\epsilon, \xi_n(y)+\epsilon) := I$ it will turn out to be enough to look at $c^n(x, y+\delta)$ only for $x \in (\xi_n(y)-\epsilon, \xi_n(y)+\epsilon)$. For such an $x$ we have

$$c^n(x, y+\delta) - c^n(x, y) \overset{(2.13)}{=} (y+\delta-x)\, K_\ell(y+\delta) - (y-x)\, K_\ell(y). \tag{2.44}$$

Let $l_2$ be the supporting tangent to $c^n(\cdot, y+\delta) - c^n(\cdot, y)$ at $\xi_n(y)$, i.e.

$$l_2(x) = c^n(\xi_n(y), y+\delta) - c^n(\xi_n(y), y) + (x - \xi_n(y))(K_\ell(y) - K_\ell(y+\delta)).$$

Hence,

$$\begin{aligned} l_1(y+\delta) + l_2(y+\delta) &\overset{(2.43)}{=} \delta(D - K_\ell(y)) \\ &\quad + c^n(\xi_n(y), y+\delta) - c^n(\xi_n(y), y) \\ &\quad + (y+\delta-\xi_n(y))(K_\ell(y) - K_\ell(y+\delta)) \\ &\overset{(2.44)}{=} \delta D \leqslant 0. \end{aligned} \tag{2.45}$$

Now, since $c^n(\cdot, y+\delta) - c^n(\cdot, y)$ is linear (and therefore convex) in the domain $I$, $l_1 + l_2$ is a supporting tangent to $c^n(\cdot, y+\delta)$ at $\xi_n(y)$, i.e.

$$(l_1 + l_2)(x) \leqslant c^n(x, y+\delta) \qquad \text{for } x \in I, \tag{2.46}$$

$$(l_1 + l_2)(\xi_n(y)) = c^n(\xi_n(y), y+\delta). \tag{2.47}$$

Recall that $\xi_n(y+\delta)$ is determined as the $x$-value where the supporting tangent to $c^n(\cdot, y+\delta)$ which passes the $x$-axis at $y+\delta$ touches $c^n(\cdot, y+\delta)$. Next we exploit the fact that $\xi_n(y+\delta) \in I$ which implies that we only need to show that $\xi_n(y+\delta) \notin (\xi_n(y)-\epsilon, \xi_n(y))$. This follows from (2.45) which yields that any supporting tangent to $c^n(\cdot, y+\delta)$ at some $\zeta \in (\xi_n(y)-\epsilon, \xi_n(y))$ must be below the $x$-axis when evaluated at $y+\delta$. We refer to [Brown et al., 2001a, Fig.7] for a graphical illustration of this fact.

Now we relax the assumption (2.40). Assume that there exists a $\delta > 0$ such that $\xi_n(y) > \xi_n(y+\delta)$. We derive a contradiction to the special case as follows. Set $y_0 := y$ and $y_n := y+\delta$. Recall that $\xi_n$ is continuous. Now we can choose $y_0 < y_1 < \cdots < y_{n-1} < y_n$ such that $\xi_n(y_0) > \xi_n(y_1) > \cdots > \xi_n(y_{n-1}) > \xi_n(y_n)$. Set $x_i := \xi_n(y_i)$, $i = 0, \ldots, n$. Observe that by monotonicity of $\xi_k$, $k < n$ the graph of $\xi_k$ intersects with at most one rectangle $(x_i, x_{i-1}) \times (y_{i-1}, y_i)$, $i = 1, \ldots, n$. Consequently, there must exist at least one integer $j$ such that the rectangle $R := (x_j, x_{j-1}) \times (y_{j-1}, y_j)$ is disjoint with the graph of every $\xi_k$, $k < n$. By construction and continuity of $y \mapsto \xi_n(y)$ $R$ is not disjoint with the graph of $\xi_n$. Inside this rectangle $R$ the conditions of the special case (2.40) are satisfied. Recalling that $\xi_n(y_j) = x_j < x_{j-1} = \xi_n(y_{j-1})$ and by continuity of $y \mapsto \xi_n(y)$, we can find two points $s_1 < s_2$ such that $z_1 = \xi_n(s_1) > \xi_n(s_2) = z_2$ and $(z_1, s_1) \in R, (z_2, s_2) \in R$. This is a contradiction. $\square$

**Lemma 2.14** (ODE for $K_n$)**.** *Let $n \in \mathbb{N}$ and let Assumption A hold. Then the mapping*

$$y \mapsto K_n(y) \tag{2.48}$$

*is absolutely continuous and non-increasing.*

*If we assume in addition that the embedding property of Theorem 2.6 is valid for the first $n-1$ marginals then for almost all $y \geqslant 0$ we have:*

*If $\xi_n(y) < y$ then*

$$K_n'(y) + \frac{K_n(y)}{y - \xi_n(y)} = K'_{\jmath_n(y)}(y) + \frac{K_{\jmath_n(y)}(y)}{y - \xi_n(y)} \tag{2.49}$$

*where $K_j'$ denotes the derivative of $K_j$ which exists for almost all $y \geqslant 0$ and $j = 1, \ldots, n$.*

*If $\xi_n(y) = y$ then*

$$K_n(y+) = K_{\jmath_n(y)}(y+). \tag{2.50}$$

*Proof.* The proof is reported in the Appendix A. □

# 3 Proof of Main Result

In this Section we prove the main result. The key step is the identification of the distribution of the maximum, cf. Proposition 3.4.

Let $n \in \mathbb{N}$. For convenience we set

$$M_0 := 0, \qquad M_i := B_{\tau_i}, \qquad i = 1, \ldots, n, \tag{3.1}$$

where $\tau_i$ is defined in Definition 2.2.

## 3.1 Basic Properties of the Embedding

Our first result shows that there is a "strong relation" between $M_n$ and $\bar{M}_n$.

**Lemma 3.1** (Relations Between $M_n$ and $\bar{M}_n$)**.** *Let $n \in \mathbb{N}$ and let Assumption A hold. Then the following implications hold.*

$$M_n \geqslant \xi_n(y) \quad \Longrightarrow \quad \bar{M}_n \geqslant y \quad \text{if } \xi_n \text{ is strictly increasing at } y. \tag{3.2}$$

*For $y \geqslant 0$ such that $\jmath_n(y) \neq 0$ we have*

$$M_{\jmath_n(y)} \geqslant \xi_n(y) > \xi_{\jmath_n(y)}(y) \qquad \Longrightarrow \qquad M_n \geqslant \xi_n(y), \tag{3.3}$$

$$\bar{M}_{\jmath_n(y)} < y, \;\; \bar{M}_n \geqslant y \qquad \Longrightarrow \qquad M_n \geqslant \xi_n(y), \tag{3.4}$$

$$\bar{M}_{\jmath_n(y)} \geqslant y, \;\; M_{\jmath_n(y)} < \xi_n(y) \qquad \Longrightarrow \qquad M_n < \xi_n(y). \tag{3.5}$$

*If $\xi_n$ is strictly increasing at $y \geqslant 0$ and $\jmath_n(y) = 0$ then the following holds.*

$$M_n \geqslant \xi_n(y) \quad \Longleftrightarrow \quad \bar{M}_n \geqslant y. \tag{3.6}$$

*Proof.* Write $\jmath = \jmath_n$. We have

$$\xi_{\jmath(y)}(y) < \xi_n(y) \leqslant \xi_i(y), \quad i = \jmath(y)+1, \ldots, n.$$

In the following we are using continuity and monotonicity of $\xi_1, \ldots, \xi_n$, cf. Lemma 2.12 and 2.39.

*Case* $\jmath(y) \neq 0$. As for implication (3.2) assume that $\xi_n$ is strictly increasing at $y$, $M_n \geqslant \xi_n(y)$ and $\bar{M}_n < y$ holds. In this case $M_n$ cannot be at the boundary $\xi_n$. It has to be at a boundary point $\xi_j(y')$ for some $j < n$ and some $y' < y$. However, this cannot be true because $\xi_n(y') < \xi_j(y')$ and hence case (2.7a) of the definition of $\tau_n$ would have been triggered.

Implication (3.3) now follows from implication (3.2) and the fact that either $M_n = M_{\jmath(y)}$ (case (2.7b)) or $M$ moves to a point at the boundary $\xi_i(y') \geqslant \xi_n(y)$ for some $i = \jmath(y)+1, \ldots, n$, $y' \geqslant y$ (case (2.7a)).

Implication (3.4) holds because the only possibility for $M$ to increase its maximum at time $\jmath(y)$, which is $< y$, to some $y' \geqslant y$ is to hit a boundary point $\xi_i(y') \geqslant \xi_n(y)$ for some $i = \jmath(y)+1, \ldots, n$.

Implication (3.5) holds because from $\bar{M}_{\jmath(y)} \geqslant y$ and $M_{\jmath(y)} < \xi_n(y)$ it follows that $M_{\jmath(y)} = \xi_i(y') < \xi_n(y) \leqslant \xi_j(y')$ for some $i \leqslant \jmath(y)$, $y' \geqslant y$, $j > \jmath(y)$. From this it follows that $M$ will stay where it is until time $n$, cf. case (2.7b).

*Case* $\jmath(y) = 0$. The condition $M_n \geqslant \xi_n(y)$ implies in a similar fashion as in (3.2) that $\bar{M}_n \geqslant y$ holds. Conversely, assume that $\bar{M}_n \geqslant y$ holds. In this case $M_n$ must be at a boundary point $\xi_i(y') \geqslant \xi_n(y)$ for some $i = 1, \ldots, n$, $y' \geqslant y$. □

As a application of Lemma 3.1 we obtain the following result.

**Lemma 3.2** (Contributions to the Maximum)**.** *Let $n \in \mathbb{N}$ and let Assumption A hold. If $\xi_n$ is strictly increasing at $y \geqslant 0$ then*

$$\mathbb{P}\left[\bar{M}_n \geqslant y\right] = \mathbb{P}\left[M_n \geqslant \xi_n(y)\right] - \mathbb{P}\left[M_{\jmath_n(y)} \geqslant \xi_n(y)\right] + \mathbb{P}\left[\bar{M}_{\jmath_n(y)} \geqslant y\right]. \tag{3.7}$$

*Proof.* Write $\jmath = \jmath_n$.

*Case* $\jmath(y) \neq 0$. Firstly, let us compute

$$\begin{aligned}
&\mathbb{P}\left[\bar{M}_n \geqslant y\right] - \mathbb{P}\left[M_n \geqslant \xi_n(y)\right] \\
\overset{(3.2)}{=}\ &\mathbb{P}\left[\bar{M}_n \geqslant y\right] - \mathbb{P}\left[M_n \geqslant \xi_n(y), \bar{M}_n \geqslant y\right] = \mathbb{P}\left[\bar{M}_n \geqslant y, M_n < \xi_n(y)\right] \\
=\ &\mathbb{P}\left[\bar{M}_n \geqslant y, M_n < \xi_n(y), \bar{M}_{\jmath(y)} \geqslant y\right] + \mathbb{P}\left[\bar{M}_n \geqslant y, M_n < \xi_n(y), \bar{M}_{\jmath(y)} < y\right] \\
\underset{(3.4)}{\overset{(3.3)}{=}}\ &\mathbb{P}\left[M_n < \xi_n(y), \bar{M}_{\jmath(y)} \geqslant y, M_{\jmath(y)} < \xi_n(y)\right].
\end{aligned}$$

Secondly, let us compute

$$\begin{aligned}
&\mathbb{P}\left[\bar{M}_{\jmath(y)} \geqslant y\right] - \mathbb{P}\left[M_{\jmath(y)} \geqslant \xi_n(y)\right] \\
=\ &\mathbb{P}\left[\bar{M}_{\jmath(y)} \geqslant y, M_{\jmath(y)} \geqslant \xi_n(y)\right] + \mathbb{P}\left[\bar{M}_{\jmath(y)} \geqslant y, M_{\jmath(y)} < \xi_n(y)\right] - \mathbb{P}\left[M_{\jmath(y)} \geqslant \xi_n(y)\right] \\
\underset{(3.3)}{\overset{(3.2)}{=}}\ &\mathbb{P}\left[\bar{M}_{\jmath(y)} \geqslant y, M_{\jmath(y)} < \xi_n(y)\right] \overset{(3.5)}{=} \mathbb{P}\left[M_n < \xi_n(y), \bar{M}_{\jmath(y)} \geqslant y, M_{\jmath(y)} < \xi_n(y)\right].
\end{aligned}$$

Comparing these two equations yields the claim.

*Case $\jmath(y) = 0$.* The claim follows from (3.6) and our convention $c_0 \equiv 0$, cf. (2.1) in Definition 2.1. $\square$

## 3.2 Law of the Maximum

Our next goal is to identify the distribution of $M_n$. We will achieve this by deriving an ODE for $\mathbb{P}\left[\bar{M}_n \geqslant \cdot\right]$ using excursion theoretical results, cf. Lemma 3.3, and link it to the ODE satisfied by $K_n$, cf. Lemma 2.14.

**Lemma 3.3** (ODE for the Maximum). *Let $n \in \mathbb{N}$ and let Assumption A hold. Then the mapping*

$$y \mapsto \mathbb{P}\left[\bar{M}_n \geqslant y\right]$$

*is absolutely continuous and for almost all $y \geqslant 0$ we have:*

*If $\xi_n(y) < y$ then*

$$\frac{\partial \mathbb{P}\left[\bar{M}_n \geqslant y\right]}{\partial y} + \frac{\mathbb{P}\left[\bar{M}_n \geqslant y\right]}{y - \xi_n(y)} = \frac{\mathbb{P}\left[\bar{M}_{\jmath_n(y)} \geqslant y\right]}{y - \xi_n(y)} + \left.\frac{\partial \mathbb{P}\left[\bar{M}_j \geqslant y\right]}{\partial y}\right|_{j=\jmath_n(y)}. \tag{3.8}$$

*If $\xi_n(y) = y$ then*

$$\mathbb{P}\left[\bar{M}_n > y\right] = \mathbb{P}\left[\bar{M}_{\jmath_n(y)} > y\right]. \tag{3.9}$$

*Proof.* Write $\jmath = \jmath_n$. We exclude all $y > 0$ which are a discontinuity of $\xi_1$. This is clearly a null-set.

The cases $n = 1, 2$ are true by [Brown et al., 2001a]. Assume by induction hypothesis that we have proven the claim for $i = 1, \ldots, n-1$.

If $\xi_n(y) = y$ then it is clear from the definition of the embedding, cf. Definition 2.2, that for any $i < n$

$$\bar{M}_n > y \qquad \Longleftrightarrow \qquad \bar{M}_i > y. \tag{3.10}$$

*Case $\jmath(y) \neq 0$.* For $\delta > 0$ we have

$$\begin{aligned} & \mathbb{P}\left[\bar{M}_n \geqslant y+\delta, \bar{M}_{\jmath(y)} < y+\delta\right] - \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_{\jmath(y)} < y\right] \\ = \; & \mathbb{P}\left[\bar{M}_n \geqslant y+\delta, \bar{M}_{\jmath(y)} < y\right] - \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_{\jmath(y)} < y\right] \\ + \; & \underbrace{\mathbb{P}\left[\bar{M}_n \geqslant y+\delta, y < \bar{M}_{\jmath(y)} < y+\delta\right]}_{=0 \text{ for } \delta > 0 \text{ small enough by definition of } \jmath(y) \text{ and continuity of } \xi_i}. \end{aligned} \tag{3.11}$$

For $r > 0$ we define

$$\begin{aligned} \bar{\xi}_j(r) &:= \max_{k: j \leqslant k \leqslant n} \left\{\xi_k(r) : \xi_k(y) = \xi_n(y)\right\}, \\ \underline{\xi}_j(r) &:= \min_{k: j \leqslant k \leqslant n} \left\{\xi_k(r) : \xi_k(y) = \xi_n(y)\right\} \end{aligned}$$

and note that

$$\bar{\xi}_{\jmath(y)}(r) \to \xi_n(y), \qquad \underline{\xi}_{\jmath(y)}(r) \to \xi_n(y) \qquad \text{as } r \to y \tag{3.12}$$

by continuity of $\xi_i$ at $y$ for $i = 1, \dots, n$.

Let $\delta > 0$. We have by excursion theoretical results, cf. e.g. [Rogers, 1989],

$$\begin{aligned}
&\mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_{\jmath(y)} < y\right] \exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \bar{\xi}_{\jmath(y)}(r)}\right) \\
&\leqslant \mathbb{P}\left[\bar{M}_n \geqslant y + \delta, \bar{M}_{\jmath(y)} < y\right] \\
&\leqslant \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_{\jmath(y)} < y\right] \exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \underline{\xi}_{\jmath(y)}(r)}\right).
\end{aligned} \tag{3.13}$$

Now we compute for $y$ such that $\xi_n(y) < y$

$$\begin{aligned}
&\frac{\mathbb{P}\left[\bar{M}_n \geqslant y + \delta, \bar{M}_{\jmath(y)} < y + \delta\right] - \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_{\jmath(y)} < y\right]}{\delta} \\
&\overset{(3.11),(3.13)}{\leqslant} \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_{\jmath(y)} < y\right] \frac{\exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \underline{\xi}_{\jmath(y)}(r)}\right) - 1}{\delta} \\
&\xrightarrow[\text{as } \delta \downarrow 0]{\text{by (3.12)}} -\frac{\mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_{\jmath(y)} < y\right]}{y - \xi_n(y)}
\end{aligned} \tag{3.14}$$

and

$$\begin{aligned}
&\frac{\mathbb{P}\left[\bar{M}_n \geqslant y + \delta, \bar{M}_{\jmath(y)} < y + \delta\right] - \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_{\jmath(y)} < y\right]}{\delta} \\
&\overset{(3.11),(3.13)}{\geqslant} \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_{\jmath(y)} < y\right] \frac{\exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \bar{\xi}_{\jmath(y)}(r)}\right) - 1}{\delta} \\
&\xrightarrow[\text{as } \delta \downarrow 0]{\text{by (3.12)}} -\frac{\mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_{\jmath(y)} < y\right]}{y - \xi_n(y)}.
\end{aligned} \tag{3.15}$$

Hence, from (3.14) and (3.15) it follows that the right-derivative of

$$y \mapsto \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_j < y\right]\Big|_{j=\jmath(y)} \tag{3.16}$$

exists. Similar arguments for $\delta < 0$ show that the left-derivative exists and is the same as the right-derivative. Local Lipschitz continuity of (3.16) then follows from (3.14) and (3.15).

Observe the obvious equality

$$\mathbb{P}\left[\bar{M}_n \geqslant y\right] = \mathbb{P}\left[\bar{M}_j \geqslant y\right] + \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_j < y\right] \tag{3.17}$$

Taking $j = \jmath(y)$ in (3.17) and fixing it, we conclude by induction hypothesis that $y \mapsto \mathbb{P}\left[\bar{M}_n > y\right]$ is locally Lipschitz continuous and hence absolutely continuous and its derivative reads

$$\frac{\partial \mathbb{P}\left[\bar{M}_n \geqslant y\right]}{\partial y} = \left.\frac{\partial \mathbb{P}\left[\bar{M}_j \geqslant y\right]}{\partial y}\right|_{j=\jmath_n(y)} + \frac{\mathbb{P}\left[\bar{M}_{\jmath_n(y)} \geqslant y\right] - \mathbb{P}\left[\bar{M}_n \geqslant y\right]}{y - \xi_n(y)}.$$

*Case* $\jmath(y) = 0$. For $\delta > 0$ we have by excursion theoretical results

$$\begin{aligned}
& \mathbb{P}\left[\bar{M}_n \geqslant y+\delta, \bar{M}_1 < y+\delta\right] - \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_1 < y\right] \\
= \; & \mathbb{P}\left[\bar{M}_n \geqslant y+\delta, \bar{M}_1 < y+\delta\right] - \mathbb{P}\left[\bar{M}_n \geqslant y+\delta, \bar{M}_1 < y\right] \\
& + \mathbb{P}\left[\bar{M}_n \geqslant y+\delta, \bar{M}_1 < y\right] - \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_1 < y\right] \\
\leqslant \; & \int_y^{y+\delta} \mathbb{P}\left[\bar{M}_1 \in \mathrm{d}s\right] \frac{\left(\xi_1(s) - \underline{\xi}_1(s)\right)^+}{s - \underline{\xi}_1(s)} \exp\left(-\int_s^{y+\delta} \frac{\mathrm{d}r}{r - \underline{\xi}_1(r)}\right) \\
& + \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_1 < y\right]\left[\exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \underline{\xi}_1(r)}\right) - 1\right].
\end{aligned} \tag{3.18}$$

Similarly, we have

$$\begin{aligned}
& \mathbb{P}\left[\bar{M}_n \geqslant y+\delta, \bar{M}_1 < y+\delta\right] - \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_1 < y\right] \\
\geqslant \; & \int_y^{y+\delta} \mathbb{P}\left[\bar{M}_1 \in \mathrm{d}s\right] \frac{\left(\xi_1(s) - \bar{\xi}_1(s)\right)^+}{s - \bar{\xi}_1(s)} \exp\left(-\int_s^{y+\delta} \frac{\mathrm{d}r}{r - \bar{\xi}_1(r)}\right) \\
& + \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_1 < y\right]\left[\exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \bar{\xi}_1(r)}\right) - 1\right].
\end{aligned} \tag{3.19}$$

From (3.18) and (3.19) it follows that the right-derivative of

$$y \mapsto \mathbb{P}\left[\bar{M}_n \geqslant y, \bar{M}_1 < y\right] \tag{3.20}$$

exists. Similar arguments for $\delta < 0$ show that the left-derivative exists and is the same as the right-derivative. Local Lipschitz continuity of (3.20) then follows from (3.18) and (3.19). Now we can conclude from (3.17)–(3.19) applied with $j = 1$ that $y \mapsto \mathbb{P}\left[\bar{M}_n \geqslant y\right]$ is locally Lipschitz continuous and hence absolutely continuous and its derivative reads

$$\begin{aligned}
& \frac{\partial \mathbb{P}\left[\bar{M}_n \geqslant y\right]}{\partial y} \\
\overset{(3.12)}{=} \; & \frac{\partial \mathbb{P}\left[\bar{M}_1 \geqslant y\right]}{\partial y} - \frac{\partial \mathbb{P}\left[\bar{M}_1 \geqslant y\right]}{\partial y} \frac{(\xi_1(y) - \xi_n(y))^+}{y - \xi_n(y)} - \frac{\mathbb{P}\left[\bar{M}_n \geqslant y\right] - \mathbb{P}\left[\bar{M}_1 \geqslant y\right]}{y - \xi_n(y)},
\end{aligned}$$

which implies by induction hypothesis

$$\frac{\mathbb{P}\left[\bar{M}_n \geqslant y\right]}{y - \xi_n(y)} + \frac{\partial \mathbb{P}\left[\bar{M}_n \geqslant y\right]}{\partial y} = 0.$$

This finishes the proof. $\square$

Finally, we argue that $\mathbb{P}\left[\bar{M}_n \geqslant y\right] = K_n(y)$ holds for all $y \geqslant 0$.

**Proposition 3.4** (Law of the Maximum)**.** *Let $n \in \mathbb{N}$ and let Assumption A hold. Assume that the embedding property of Theorem 2.6 is valid for the first $n-1$ marginals. Then for all $y \geqslant 0$ we have*

$$\mathbb{P}\left[\bar{M}_n \geqslant y\right] = K_n(y). \tag{3.21}$$

*Proof.* The case $n = 1$ holds by the Azéma-Yor embedding. Assume by induction hypothesis that

$$K_i = \mathbb{P}\left[\bar{M}_i \geqslant \cdot\right], \qquad i = 1, \ldots, n-1.$$

In Lemma 2.14 and 3.3 we derived an ODE for $K_n$ and $\mathbb{P}\left[\bar{M}_n \geqslant \cdot\right]$, respectively, in terms of $K_1, \ldots, K_{n-1}$ and $\mathbb{P}\left[\bar{M}_1 \geqslant \cdot\right], \ldots, \mathbb{P}\left[\bar{M}_{n-1} \geqslant \cdot\right]$, respectively. These ODEs are valid for a.e. $y \geqslant 0$. By induction hypothesis both drivers of these ODEs coincide everywhere and hence the claim follows from the boundary conditions

$$\begin{aligned} K_n(y) &\to 0 \quad \text{as } y \to \infty, & K_n(y) &\to 1 \quad \text{as } y \to 0, \\ \mathbb{P}\left[\bar{M}_n \geqslant y\right] &\to 0 \quad \text{as } y \to \infty, & \mathbb{P}\left[\bar{M}_n \geqslant y\right] &\to 1 \quad \text{as } y \to 0, \end{aligned}$$

absolute continuity of $K_n$ and $\mathbb{P}\left[\bar{M}_n \geqslant \cdot\right]$ and the fact that the ODE

$$\Big(\mathbb{P}\left[\bar{M}_n \geqslant y\right] - K_n(y)\Big)' = -\frac{\mathbb{P}\left[\bar{M}_n \geqslant y\right] - K_n(y)}{y - \xi_n(y)}, \qquad \mathbb{P}\left[\bar{M}_n \geqslant 0\right] - K_n(0) = 0,$$

has unique solution given by 0. $\square$

## 3.3 Embedding Property

In this subsection we prove that the stopping times $\tau_1, \ldots, \tau_n$ from Definition 2.2 embed the laws $\mu_1, \ldots, \mu_n$ if Assumption A is in place. More precisely, given Proposition 3.4 above and by inductive reasoning, to complete the proof of Theorem 2.6 we only need to show the following:

**Proposition 3.5** (Embedding)**.** *In the setup of Theorem 2.6 we have*

$$B_{\tau_n} \sim \mu_n \tag{3.22}$$

*and $(B_{\tau_n \wedge t})_{t \geqslant 0}$ is a uniformly integrable martingale.*

*Proof.* The case $n = 1$ is just the Azéma-Yor embedding. By induction hypothesis, assume that the claim holds for all $i \leqslant n - 1$.

We claim that $\xi_n$ ranges continuously over the full support of $\mu_n$. This is because, firstly, we know from Lemma 2.13 that $\xi_2, \dots, \xi_n$ are continuous. Secondly, we have by using $l_{\mu_n} \leqslant l_{\mu_i}$ that

$$\inf_{\zeta \leqslant 0} \frac{c^n(\zeta,0)}{0-\zeta} \geqslant \inf_{\zeta \leqslant 0} \min_{1 \leqslant i < n} \left\{ \underbrace{\frac{c_n(\zeta)-c_i(\zeta)}{0-\zeta}}_{\geqslant 0} + \underbrace{K_i(0)}_{=1} \right\} \wedge \underbrace{\frac{c_n(l_{\mu_n})}{0-l_{\mu_n}}}_{=1} = 1$$

which shows that $\xi_n(0) = l_{\mu_n}$. Furthermore, by using $r_{\mu_n} \geqslant r_{\mu_i}$ we have from (2.11) and (2.16) that

$$\xi_n(r_{\mu_n}) = r_{\mu_n}.$$

Let $y > 0$ be such that $\xi_n$ is differentiable and strictly increasing at $y$, $\xi_n(y)$ is not an atom of neither $\mu_n$ nor $\mu_{\jmath_n(y)}$ and $y$ is not a discontinuity of $\xi_1$. Note that for such a $y$ equation (A.6) holds because of (A.7). Applying previous results we obtain

$$\begin{aligned}
&\mathbb{P}\left[M_n \geqslant \xi_n(y)\right] - \mathbb{P}\left[M_{\jmath_n(y)} \geqslant \xi_n(y)\right] + \mathbb{P}\left[\bar{M}_{\jmath_n(y)} \geqslant y\right] \\
\overset{\text{Lemma 3.2}}{=} \; &\mathbb{P}\left[\bar{M}_n \geqslant y\right] \\
\overset{\text{Prop. 3.4}}{=} \; &K_n(y) \\
\overset{\text{(A.6)}}{=} \; &-c_n'(\xi_n(y)) + c_{\jmath_n(y)}'(\xi_n(y)) + K_{\jmath_n(y)}(y),
\end{aligned}$$

which implies by induction hypothesis that

$$\mathbb{P}\left[M_n \geqslant \xi_n(y)\right] = -c_n'(\xi_n(y)) = \mu_n([\xi_n(y), \infty)).$$

We have matched the distribution of $M_n$ to $\mu_n$ at almost all points inside the support. The embedding property follows.

Now we prove uniform integrability by applying a result from [Azéma et al., 1980] which states that if

$$\lim_{x \to \infty} x \mathbb{P}\left[|\bar{B}|_{\tau_n} \geqslant x\right] = 0 \tag{3.23}$$

then $(B_{\tau_n \wedge t})_{t \geqslant 0}$ is uniformly integrable.

Let us verify (3.23). Set $H_x = \inf\{t > 0 : B_t = x\}$. We have (here $\xi_i^{-1}$ denotes the left-continuous inverse of $\xi_i$)

$$\begin{aligned}
\mathbb{P}\left[|\bar{B}|_{\tau_n} \geqslant x\right] &\leqslant \mathbb{P}\left[H_{-x} < H_{\max_{i \leqslant n} \xi_i^{-1}(-x)}\right] + \mathbb{P}\left[\bar{B}_{\tau_n} \geqslant x\right] \\
&= \frac{\max_{i \leqslant n} \xi_i^{-1}(-x)}{x + \max_{i \leqslant n} \xi_i^{-1}(-x)} + K_n(x).
\end{aligned}$$

From the definition of $\xi_n$, cf. (2.4), and the properties of $b_i$, cf. (2.11) we have

$$\max_{i\leqslant n} \xi_i^{-1}(-x) \leqslant \max_{i\leqslant n} b_i(-x) \underset{x\to\infty}{\longrightarrow} 0$$

and hence, recalling the definition of $\mu_n^{\mathrm{HL}}$ in (2.23),

$$\lim_{x\to\infty} x\mathbb{P}\left[|\bar{B}|_{\tau_n} \geqslant x\right] \leqslant \lim_{x\to\infty} xK_n(x) \leqslant \lim_{x\to\infty} x\frac{c_n(b_n^{-1}(x))}{x-b_n^{-1}(x)} = \lim_{x\to\infty} x\mu_n^{\mathrm{HL}}([x,\infty)) = 0.$$

This finishes the proof. □

# 4 Discussion of Assumption A and Extensions

In this section we present a simple example of probability measures $\mu_1, \mu_2, \mu_3$ which violate Assumption A(ii) and where the stopping boundaries $\xi_1, \xi_2, \xi_3$, obtained via (2.4), fail to embed $(\mu_1, \mu_2, \mu_3)$.

Motivated by this counterexample we sketch how an iterated Azéma-Yor type embedding in the three marginal case $n = 3$ looks like, i.e. in the absence of Assumption A(ii). This will lead us to redefine $\xi_3$ in certain regions. More precisely, we show how to modify the optimization problem from which $\xi_3$ is determined in order to obtain the embedding property. We will also point out some major differences of this (general) embedding to the embedding in the presence of Assumption A(ii).

## 4.1 Counterexample for Assumption A(ii)

In Figure 4.1 we define measures via their potentials

$$U\mu : \mathbb{R} \to \mathbb{R}, \quad U\mu(x) := -\int_{\mathbb{R}} |u-x|\,\mu(\mathrm{d}u). \tag{4.1}$$

We refer to [Obłój, 2004, Proposition 2.3] for some properties of $U\mu$.

The measures are given as

$$\mu_1(-1) = \frac{2}{3}, \quad \mu_1(2) = \frac{1}{3}, \tag{4.2}$$

$$\mu_2(-3) = \frac{2}{7}, \quad \mu_2\left(\frac{1}{2}\right) = \frac{18}{35}, \quad \mu_2(3) = \frac{1}{5}, \tag{4.3}$$

$$\mu_3(-3) = \frac{2}{7}, \quad \mu_3(-2) = \frac{9}{35}, \quad \mu_3(3) = \frac{16}{35}. \tag{4.4}$$

Observe that the embedding for $(\mu_1, \mu_2, \mu_3)$ is unique: We write $H_{a,b}$ for the exit time of $[a, b]$ and denote $H_{a,b} \circ \theta_\tau := \inf\{t > \tau : B_t \notin (a, b)\}$. Then the

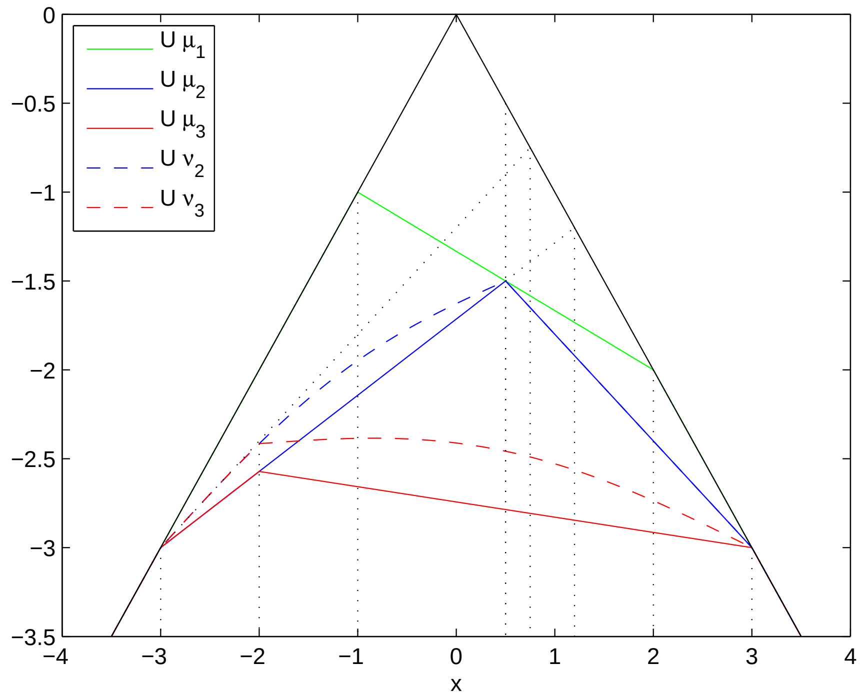


Figure 4.1: Potentials of $\mu_1, \mu_2, \mu_3, \nu_2$ and $\nu_3$.

embedding $(\tau_1, \tau_2, \tau_3)$ can be written as

$$\tau_1 = H_{-1,2}, \quad \tau_2 = H_{-3,\frac{1}{2}} \circ \theta_{\tau_1}\Big|_{B_{\tau_1}=-1} + H_{\frac{1}{2},3} \circ \theta_{\tau_1}\Big|_{B_{\tau_1}=2}, \quad \tau_3 = H_{-2,3} \circ \theta_{\tau_2}. \tag{4.5}$$

As mentioned earlier, our construction yields the same first two stopping boundaries as the method of [Brown et al., 2001a]. In this case, cf. Figure 4.2,

$$\xi_1(y) := \begin{cases} -1 & \text{if } y \in [0,2), \\ y & \text{else,} \end{cases} \qquad \xi_2(y) := \begin{cases} -3 & \text{if } y \in [0,\frac{1}{2}), \\ \frac{1}{2} & \text{if } y \in [\frac{1}{2},3), \\ y & \text{else.} \end{cases}$$

This already shows that our embedding fails to embed $\mu_2$: while the embedding for $\mu_1$ is valid by the Azéma-Yor type embedding, the embedding for $\mu_2$ is invalid. To see this one just has to compare the stopping boundary $\xi_2$ in the Definition of $\tau_2$ with (4.5). In Section (4.2) we will recall from [Brown et al., 2001a] how the stopping time $\tau_2$ has to be modified.

However, the point we want to make, is that the embedding for $\mu_3$ fails and cannot be achieved by using $\xi_3$: the optimization problem (2.4) does not

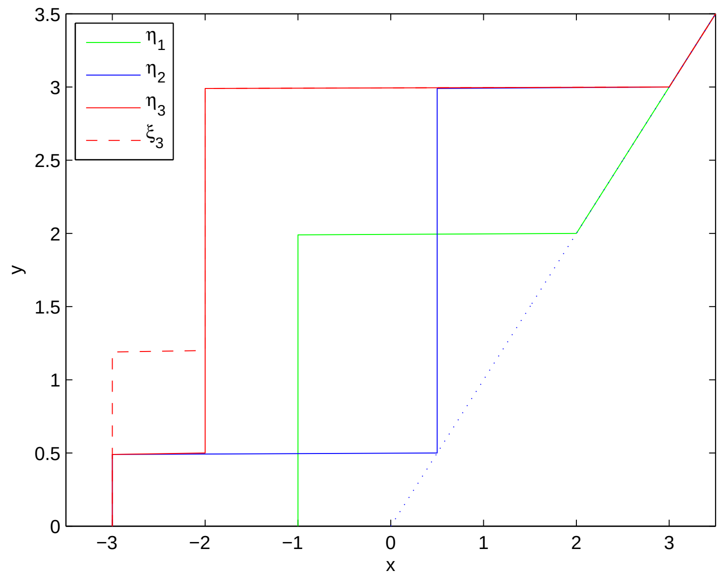


Figure 4.2: We illustrate the (unique) boundaries $\eta_1, \eta_2, \eta_3$ required for the embedding of $(\mu_1, \mu_2, \mu_3)$ from (4.2)–(4.4) and the stopping boundary $\xi_3$ obtained from (2.4). We already know that $\eta_1 = \xi_1$ and $\eta_2 = \xi_2$. In order to ensure the embedding for $\mu_2$, the mass in the region $(-1) \times (\frac{1}{2}, 2)$ is redistributed to $(-3) \times (\frac{1}{2}, 2)$ and $(0.5) \times (\frac{1}{2}, 2)$. Note that the case $\eta_2(y) = y = \xi_2(y)$ is possible and required to define the embedding. After the embedding for $(\mu_1, \mu_2)$, the [Brown et al., 2001a] embedding, our goal is to define the embedding for $\mu_3$ on top of that "in the same fashion". In this example $\eta_3$ is implied directly by (4.5). In Section 4.2 we show how to obtain it in general.

return the third (unique) stopping boundary which is required for the embedding of $(\mu_1, \mu_2, \mu_3)$. Indeed, for sufficiently small $y > \frac{1}{2}$, in the region $\zeta < \min(\xi_1(y), \xi_2(y)) = -1$ we are looking at the minimization of $\zeta \mapsto \frac{c_3(\zeta)}{y-\zeta}$ which is attained by $\xi_3(y) = -3 < -2$ since $\mu_3$ has an atom at $-3$. Consequently, there will be a positive probability to hit $-3$ after $\tau_2$. This contradicts (4.4).

This example does not contradict our main result because Assumption $A$(ii)(a) is not satisfied for $i = 2$ and $y = \frac{1}{2}$, where $\zeta = -3$ minimizes the objective function but $c_2\left(\frac{1}{2}\right) = c_1\left(\frac{1}{2}\right)$ holds.

## 4.2 Sketch for General Embedding in the Case $n = 3$

In the example of the measures $(\mu_1, \mu_2, \mu_3)$ from (4.2)–(4.4), we note however that the (unique) embedding resembles something which might be termed "iterated Azéma-Yor type embedding" although this embedding does not satisfy the rela-

tions from Lemma 3.1. Consequently, one might conjecture that a modification of the optimization problem (2.4) and a relaxation of Lemma 3.1 might fix this problem. We now explain in which sense this is true. The purpose of this section is to outline the main arguments and hence we sacrifice occasionally on rigor.

In order to understand the problem in more detail, we need to recall from [Brown et al., 2001a] how the embedding for $\mu_2$ looks like in general. It reads

$$\tau_2^{\mathrm{BHR}} := \begin{cases} \tau_2' & \text{if } \xi_2^{-1} \neq \xi_1(\bar{B}_{\tau_1}) \text{ and } \xi_1(\bar{B}_{\tau_1}) < \xi_2(\bar{B}_{\tau_1}), \\ \tau_2 & \text{else}, \end{cases} \tag{4.6}$$

where $\tau_2'$ is some stopping time embedding whose existence is established in [Brown et al., 2001a] and corresponds to the case where $\mu_1$ and $\mu_2$ have intervals which have same mass, mean and are in convex order. One important property of $\tau_2'$ is that it does not change the current maximum, i.e. $\bar{B}_{\tau_1} = \bar{B}_{\tau_2'}$. In general there will be infinitely many such stopping times $\tau_2'$. Although this is not true for $(\mu_1, \mu_2, \mu_3)$ because their embedding is unique, we illustrate this claim by the measures $(\mu_1, \nu_2, \nu_3)$ which are defined via their potentials in Figure 4.1. This example also suggests that the atomic nature of the measures (which results in discontinuous stopping boundaries) is not the main problem but rather the intervals with same mass, mean and convex order.

Let $\xi_1$ and $\xi_2$ be defined as in (2.4) but now we let $\tau_2 := \tau_2^{\mathrm{BHR}}$ be the [Brown et al., 2001a] embedding, cf. (4.6). Also set $M_2 = B_{\tau_2}$.

Now our goal is to define an embedding $\tilde{\tau}_3$ for the third marginal on top of the [Brown et al., 2001a] embedding in a situation as in $(\mu_1, \nu_2, \nu_3)$. We still want to define our iterated Azéma-Yor type embedding through a stopping rule based on some stopping boundary $\tilde{\xi}_3$ as a first exit time,

$$\tilde{\tau}_3 := \begin{cases} \inf\left\{t \geqslant \tau_2^{\mathrm{BHR}} : B_t \leqslant \tilde{\xi}_3(\bar{B}_t)\right\} & \text{if } B_{\tau_2^{\mathrm{BHR}}} > \tilde{\xi}_3(\bar{B}_{\tau_2^{\mathrm{BHR}}}), \\ \tau_2^{\mathrm{BHR}} & \text{else}, \end{cases} \tag{4.7}$$

and prove that this is a valid embedding of $\mu_3$. We observe that now the choice of $\tau_2'$ in the definition of $\tau_2 = \tau_2^{\mathrm{BHR}}$ matters both for the embedding and the distribution of the maximum $\bar{B}_{\tilde{\tau}_3}$. Similarly as in [Brown et al., 2001a] we expect that this will be only possible if the procedure which produces $\tilde{\xi}_3$ yields a continuous $\tilde{\xi}_3$. We restrict to this case because otherwise an additional step is likely to be required.

With this, a more canonical approach in the context of Lemma 3.1 is to write

$$\mathbb{P}\left[\bar{M}_3 \geqslant y\right] = \mathbb{P}\left[M_3 \geqslant \tilde{\xi}_3(y)\right] + \text{“error-term”},$$

which we formalize in (4.19). As it will turn out, this “error-term” provides a suitable “book-keeping procedure” to keep track of the masses in the embedding. Then we proceed along the lines of the proof of our main result.

As mentioned above, we do not present a rigorous proof for the embedding property but rather outline the main ideas. For simplicity, we also assume that $\xi_2$ has only one discontinuity, i.e. $\Xi_- := \xi_2(\underline{y}-) < \xi_2(\underline{y}+) := \Xi_+$ for some $\underline{y} \geqslant 0$. Denote $\bar{y} := \xi_1^{-1}(\Xi_+)$.

### 4.2.1 Redefining $\xi_3$ and $K_3$

Define the following auxiliary terms,

$$F(\zeta, y; \tau_2') \quad := \mathbb{1}_{\left\{\bar{M}_1 \in [y,\bar{y}], y \geqslant \underline{y}\right\}} \left(\zeta - M_2\right)^+, \tag{4.8}$$

$$f^{\mathrm{iAY}}(\zeta, y; \tau_2') := \mathbb{E}\left[F(\zeta, y; \tau_2')\right]. \tag{4.9}$$

Note that for $y \geqslant \underline{y}$,

$$\frac{\partial f^{\mathrm{iAY}}}{\partial \zeta}(\zeta, y; \tau_2') = \mathbb{P}\left[\bar{M}_1 \in [y, \bar{y}], M_2 < \zeta\right], \tag{4.10}$$

and

$$\begin{aligned} \frac{\partial f^{\mathrm{iAY}}}{\partial y}(\zeta, y; \tau_2') &= -\mathbb{E}\left[\frac{\mathbb{1}_{\left\{\bar{M}_1 \in \mathrm{d}y, M_2 < \zeta\right\}}}{\mathrm{d}y}\right]\zeta + \mathbb{E}\left[\frac{\mathbb{1}_{\left\{\bar{M}_1 \in \mathrm{d}y, M_2 < \zeta\right\}}}{\mathrm{d}y} M_2\right] \\ &= -\Big(\zeta - \alpha(\zeta, y; \tau_2')\Big)\frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y, M_2 < \zeta\right]}{\mathrm{d}y} \end{aligned} \tag{4.11}$$

where

$$\alpha(\zeta, y; \tau_2') := \mathbb{E}\left[M_2 \middle| \bar{M}_1 = y, M_2 < \zeta\right], \tag{4.12}$$

$$\beta(\zeta, y; \tau_2') := \mathbb{E}\left[M_2 \middle| \bar{M}_1 = y, M_2 \geqslant \zeta\right]. \tag{4.13}$$

With these definitions we have by the properties of $\tau_2'$,

$$\begin{aligned} &\alpha(\zeta, y; \tau_2')\frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y, M_2 < \zeta\right]}{\mathrm{d}y} + \beta(\zeta, y; \tau_2')\frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y, M_2 \geqslant \zeta\right]}{\mathrm{d}y} \\ =&\xi_1(y)\frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y\right]}{\mathrm{d}y}. \end{aligned} \tag{4.14}$$

We now redefine $\xi_3$ and $K_3$ from (2.4) and (2.6), respectively, and denote the new definition by $\tilde{\xi}_3$ and $\tilde{K}_3$. To this end, introduce the function

$$\tilde{c}^3(\zeta, y) := \begin{cases} c_3(\zeta) - f^{\mathrm{iAY}}(\zeta, y; \tau_2') & \text{if } f^{\mathrm{iAY}}(\zeta, y; \tau_2') > 0 \text{ and } \zeta < \Xi_+, \quad (4.15a) \\ c^3(\zeta, y) & \text{else.} \quad (4.15b) \end{cases}$$

Now set

$$\tilde{\xi}_3(y) := \arg\min_{\zeta < y} \frac{\tilde{c}^3(\zeta, y)}{y - \zeta} \tag{4.16}$$

and

$$\tilde{K}_3(y) := \frac{\tilde{c}^3(\tilde{\xi}_3(y), y)}{y - \tilde{\xi}_3(y)}. \tag{4.17}$$

The first order condition for optimality of $\tilde{\xi}_3(y)$ reads

$$\tilde{K}_3(y) + c_3'(\tilde{\xi}_3(y)) - \frac{\partial f^{\mathrm{iAY}}}{\partial \zeta}(\tilde{\xi}_3(y), y; \tau_2') = 0. \tag{4.18}$$

We stress that now the quantities $\tilde{\xi}_3$ and $\tilde{K}_3$ depend on $\tau_2'$.

### 4.2.2 Embedding Property

In the following we assume that $\zeta$ and $y$ are such that $f^{\mathrm{iAY}}(\zeta, y; \tau_2') > 0$ and $\zeta < \Xi_+$. If this is not true the arguments from Sections 2 and 3 apply.

By definition of the embedding in (4.7) we have, assuming $\tilde{\xi}_3$ is non-decreasing,

$$\begin{aligned}
\mathbb{P}\left[\bar{M}_3 \geqslant y\right] &= \mathbb{P}\left[M_3 \geqslant \tilde{\xi}_3(y)\right] + \mathbb{P}\left[\bar{M}_3 \geqslant y, M_3 < \tilde{\xi}_3(y)\right] \\
&= \mathbb{P}\left[M_3 \geqslant \tilde{\xi}_3(y)\right] + \mathbb{P}\left[\bar{M}_1 \geqslant y, M_2 < \tilde{\xi}_3(y)\right] \\
&= \mathbb{P}\left[M_3 \geqslant \tilde{\xi}_3(y)\right] + \mathbb{P}\left[\bar{M}_1 \in [y, \bar{y}], M_2 < \tilde{\xi}_3(y)\right] \\
&\overset{(4.10)}{=} \mathbb{P}\left[M_3 \geqslant \tilde{\xi}_3(y)\right] + \frac{\partial f^{\mathrm{iAY}}}{\partial \zeta}(\tilde{\xi}_3(y), y; \tau_2').
\end{aligned} \tag{4.19}$$

We observe the obvious equation

$$\mathbb{P}\left[\bar{M}_3 \geqslant y\right] = \mathbb{P}\left[\bar{M}_1 < \underline{y}, M_3 \geqslant y\right] + \mathbb{P}\left[\bar{M}_1 \in [\underline{y}, y), \bar{M}_3 \geqslant y\right] + \mathbb{P}\left[\bar{M}_1 \geqslant y\right] \tag{4.20}$$

and note that

$$\begin{aligned}
&\left.\frac{\partial \mathbb{P}\left[\bar{M}_1 \in [\underline{y}, y), \bar{M}_3 \geqslant m\right]}{\partial y}\right|_{m=y} = p(\tilde{\xi}_3(y), y; \tau_2') := \\
&\frac{\mathbb{P}\left[M_2 > \tilde{\xi}_3(y), \bar{M}_1 \in \mathrm{d}y\right]}{\mathrm{d}y} \cdot \mathbb{P}\left[H_y \circ \theta_{\tau_2'} < H_{\tilde{\xi}_3(y)} \circ \theta_{\tau_2'} \middle| M_2 > \tilde{\xi}_3(y), \bar{M}_1 = y\right]
\end{aligned} \tag{4.21}$$

where as above $H_x := \inf\{t \geqslant 0 : B_t = x\}$.

Hence, by (4.20), (4.21) and similar excursion theoretical results as in the proof of Lemma 3.3,

$$\begin{aligned}\frac{\partial}{\partial y}\mathbb{P}\left[\bar{M}_3 \geqslant y\right] &= -\frac{\mathbb{P}\left[\bar{M}_3 \geqslant y\right]}{y-\tilde{\xi}_3(y)} + p(\tilde{\xi}_3(y), y; \tau_2') + \frac{\partial \mathbb{P}\left[\bar{M}_1 \geqslant y\right]}{\partial y} + \frac{\mathbb{P}\left[\bar{M}_1 \geqslant y\right]}{y-\tilde{\xi}_3(y)} \\ &\overset{(3.8)}{=} -\frac{\mathbb{P}\left[\bar{M}_3 \geqslant y\right]}{y-\tilde{\xi}_3(y)} + p(\tilde{\xi}_3(y), y; \tau_2') - \frac{\tilde{\xi}_3(y)-\xi_1(y)}{y-\tilde{\xi}_3(y)}\frac{\partial \mathbb{P}\left[\bar{M}_1 \geqslant y\right]}{\partial y}.\end{aligned} \tag{4.22}$$

Hence, by similar calculations as in (A.17),

$$\begin{aligned}\tilde{K}_3'(y) &\overset{(4.18)}{=} -\frac{\tilde{K}_3(y)}{y-\tilde{\xi}_3(y)} - \frac{\frac{\partial f^{\mathrm{iAY}}}{\partial y}(\tilde{\xi}_3(y), y; \tau_2')}{y-\tilde{\xi}_3(y)} \\ &\overset{(4.11)}{=} -\frac{\tilde{K}_3(y)}{y-\tilde{\xi}_3(y)} + \frac{\tilde{\xi}_3(y)-\alpha(\tilde{\xi}_3(y), y)}{y-\tilde{\xi}_3(y)}\frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y, M_2 < \tilde{\xi}_3(y)\right]}{\mathrm{d}y} \\ &\overset{(4.14)}{=} -\frac{\tilde{K}_3(y)}{y-\tilde{\xi}_3(y)} + \frac{\tilde{\xi}_3(y)-\xi_1(y)}{y-\tilde{\xi}_3(y)}\frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y\right]}{\mathrm{d}y} + \frac{\beta(y)-\tilde{\xi}_3(y)}{y-\tilde{\xi}_3(y)}\frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y, M_2 \geqslant \tilde{\xi}_3(y)\right]}{\mathrm{d}y} \\ &\overset{(4.21)}{=} -\frac{\tilde{K}_3(y)}{y-\tilde{\xi}_3(y)} - \frac{\tilde{\xi}_3(y)-\xi_1(y)}{y-\tilde{\xi}_3(y)}\frac{\partial \mathbb{P}\left[\bar{M}_1 \geqslant y\right]}{\partial y} + p(\tilde{\xi}_3(y), y; \tau_2').\end{aligned} \tag{4.23}$$

Consequently, as in and together with Proposition 3.5, for all $y$,

$$\tilde{K}_3(y) = \mathbb{P}\left[\bar{M}_3 \geqslant y\right] \tag{4.24}$$

and

$$-c_3'(\tilde{\xi}_3(y)) = \mathbb{P}\left[M_3 \geqslant \tilde{\xi}_3(y)\right]. \tag{4.25}$$

If $\tilde{\xi}_3$ happens to be discontinuous, then we cannot guarantee the embedding property for $\mu_3$. However, we expect that similar arguments as those in [Brown et al., 2001a, Section 3.5] will yield a stopping time $\tau_3'$ which does the job for the jump-intervals of $\tilde{\xi}_3$. We do not pursue these matters here.

# A Appendix: Proof of Lemma 2.14

In order to prove Lemma 2.14 we require to prove, inductively, several auxiliary results along the way. We now state and prove a Lemma which contains the statement of Lemma 2.14.

**Lemma A.1.** *Let $n \in \mathbb{N}$ and let Assumption A hold. Then the mapping*

$$y \mapsto K_n(y) \tag{A.1}$$

*is absolutely continuous and non-increasing.*

*If we assume in addition that the embedding property of Theorem 2.6 is valid for the first $n-1$ marginals then for almost all $y \geqslant 0$ we have:*

*If $\xi_n(y) < y$ then*

$$K_n'(y) + \frac{K_n(y)}{y - \xi_n(y)} = K_{\jmath_n(y)}'(y) + \frac{K_{\jmath_n(y)}(y)}{y - \xi_n(y)} \tag{A.2}$$

*where $K_j'$ denotes the derivative of $K_j$ which exists for almost all $y \geqslant 0$ and $j = 1, \dots, n$.*

*If $\xi_n(y) = y$ then*

$$K_n(y+) = K_{\jmath_n(y)}(y+). \tag{A.3}$$

*For $x > 0$ the mapping*

$$c^n : (x, \infty) \to \mathbb{R}, \quad y \mapsto c^n(x, y) \tag{A.4}$$

*is locally Lipschitz continuous, non-decreasing and for almost all $y > 0$*

$$\frac{\partial c^n}{\partial y}(x, y)\bigg|_{x=\xi_n(y)} = K_{\jmath_n(y)}(y) + (y - \xi_n(y))K_{\jmath_n(y)}'(y). \tag{A.5}$$

*If the mapping $c^n(\cdot, y)$ is differentiable at $\xi_n(y)$ and $\xi_n'(y) > 0$ then for almost all $y \geqslant 0$*

$$K_n(y) + c_n'(\xi_n(y)) - c_j'(\xi_n(y)) - K_j(y) = 0 \tag{A.6}$$

*for $j = \jmath_n(y)$ and $j$ such that $n > j > \jmath_n(y)$ and $\xi_n(y) = \xi_j(y)$. In the case of non-smoothness of $c^n(\cdot, y)$ at $\xi_n(y)$ we have*

$$\xi_n'(y) = 0 \tag{A.7}$$

*for $y$ such that the slope of the supporting tangent to $c^n(\cdot, y)$ at $\xi_n(y)$ which crosses the $x$-axis at $y$ does not equal the right-derivative of $c^n(\cdot, y)$ at $\xi_n(y)$.*

*Proof.* We prove the claim by induction over $n$. The induction basis $n = 1$ holds by definition and [Brown et al., 2001a, Lemma 2.6].

Now assume that the claim holds for all $i = 1, \dots, n-1$.

*Induction step for $c^n$.* We have

$$\begin{aligned} c^n(x, y+\delta) - c^n(x, y) = &- \left[c_{\imath_n(x;y+\delta)}(x) - (y + \delta - x)K_{\imath_n(x;y+\delta)}(y+\delta)\right] \\ &+ \left[c_{\imath_n(x;y)}(x) - (y - x)K_{\imath_n(x;y)}(y)\right]. \end{aligned} \tag{A.8}$$

Firstly, consider the case when there exists a $\delta' > 0$ such that for all $|\delta| < \delta'$ we have $\imath_n(x; y) = \imath_n(x; y + \delta)$. Equation (A.8) simplifies and we have

$$\begin{aligned} & c^n(x, y+\delta) - c^n(x, y) = (y+\delta-x)K_{\imath_n(x;y)}(y+\delta) - (y-x)K_{\imath_n(x;y)}(y) \\ = \;& (y+\delta-\xi_{\imath_n(x;y)}(y))K_{\imath_n(x;y)}(y+\delta) - (y-\xi_{\imath_n(x;y)}(y))K_{\imath_n(x;y)}(y) \\ & +(x-\xi_{\imath_n(x;y)}(y))\left[K_{\imath_n(x;y)}(y) - K_{\imath_n(x;y)}(y+\delta)\right] \\ \overset{(2.14)}{\leqslant} \;& c^{\imath_n(x;y)}(\xi_{\imath_n(x;y)}(y), y+\delta) - c^{\imath_n(x;y)}(\xi_{\imath_n(x;y)}(y), y) \\ & +(y-\xi_{\imath_n(x;y)}(y))\left[K_{\imath_n(x;y)}(y) - K_{\imath_n(x;y)}(y+\delta)\right] \\ \leqslant \;& \max_{i<n}\left\{c^i(\xi_i(y), y+\delta) - c^i(\xi_i(y), y) + (y-\xi_i(y))\left[K_i(y) - K_i(y+\delta)\right]\right\} \\ \leqslant \;& \mathrm{const}(y)\cdot|\delta| \end{aligned} \tag{A.9}$$

by induction hypothesis and where const$(y)$ denotes a constant depending on $y$. A similar computation shows for $|\delta|$ small enough

$$\begin{aligned} & c^n(x, y) - c^n(x, y+\delta) \\ \overset{(2.14)}{\leqslant} \;& c^{\imath_n(x;y)}(\xi_{\imath_n(x;y)}(y+\delta), y) - c^{\imath_n(x;y)}(\xi_{\imath_n(x;y)}(y+\delta), y+\delta) \\ & +(x-\xi_{\imath_n(x;y)}(y+\delta))\left[K_{\imath_n(x;y)}(y+\delta) - K_{\imath_n(x;y)}(y)\right] \\ \leqslant \;& \begin{cases} \mathrm{const}(y)\cdot|\delta| & \text{if } \delta < 0, \\ 0 & \text{if } \delta \geqslant 0. \end{cases} \end{aligned} \tag{A.10}$$

Monotonicity of $c^n(x, \cdot)$ follows. Equation (A.10) together with (A.9) imply the local Lipschitz continuity. Plugging $x = \xi_n(y)$ into (A.8), direct computation shows that (A.5) holds.

Secondly, consider the case when $\imath_n(x; \cdot)$ jumps at $y$. Note that this is only possible when $x$ satisfies

$$x = \xi_{\imath_n(x;y-\delta)}(y) \qquad \text{for } \delta > 0 \text{ small enough,} \tag{A.11}$$

i.e. when $x = \xi_k(y)$ for the index $k = \imath_n(x; y-\delta) > \jmath_n(y)$. By (2.9) there exists a $\delta' > 0$ such that $\imath_n(x; y+\delta) = \imath_n(x; y)$ for all $0 \leqslant \delta < \delta'$. Hence, for $\delta > 0$ small enough, $|c^n(x, y+\delta) - c^n(x, y)|$ has the same upper bound as in the first case. Monotonicity of $c^n(x, \cdot)$ follows.

Furthermore, for $\delta > 0$ we have for the $x$ in (A.11) that $\imath_n(x; y-\delta) > \imath_n(x; y)$ holds. For notational simplicity we only consider the case

$$\imath_{\imath_n(x;y-\delta)}(x; y-\delta) = \imath_n(x; y).$$

The general case follows by the same arguments. We deduce from (A.8) the following two equations,

$$\begin{aligned} c^n(x, y-\delta) - c^n(x, y) &\overset{(2.14)}{\leqslant} (y-\delta-x)K_{\imath_{\imath_n(x;y-\delta)}(x;y-\delta)}(y-\delta) - (y-x)K_{\imath_n(x;y)}(y) \\ &= (y-\delta-x)K_{\imath_n(x;y)}(y-\delta) - (y-x)K_{\imath_n(x;y)}(y) \end{aligned} \tag{A.12}$$

and

$$c^n(x,y) - c^n(x,y-\delta) \overset{\text{(A.11)}}{=} -(y-\delta-x)K_{\imath_n(x;y-\delta)}(y-\delta) + (y-x)K_{\imath_n(x;y-\delta)}(y).$$

Now the local Lipschitz continuity of $c^n(x,\cdot)$ follow from the above two equations and the first case.

We prove (A.5) by computing the required right- and left-derivative of $c^n(x,\cdot)$ at $x = \xi_n(y)$. The right-derivative is simply, using (2.9) and (A.8),

$$K_{\jmath_n(y)}(y) + (y - \xi_n(y))K'_{\jmath_n(y)}(y) \tag{A.13}$$

and the left-derivative is, writing $k = \imath_n(\xi_n(y); y-) > \imath_n(\xi_n(y); y) = \jmath_n(y)$,

$$\begin{aligned}
&\lim_{\delta\uparrow 0}\frac{1}{\delta}\Big( - c_k(\xi_n(y)) + (y+\delta-\xi_n(y))K_k(y+\delta) \\
&\qquad\qquad + c_{\jmath_n(y)}(\xi_n(y)) - (y-\xi_n(y))K_{\jmath_n(y)}(y)\Big) \\
\overset{\xi_n(y)=\xi_k(y)}{=}\; &\lim_{\delta\uparrow 0}\frac{1}{\delta}\Big((y+\delta-\xi_n(y))K_k(y+\delta) - (y-\xi_n(y))K_k(y)\Big) \\
= \;& K_k(y) + (y-\xi_n(y))K'_k(y) \overset{\text{(A.2)}}{=} K_{\jmath_n(y)}(y) + (y-\xi_n(y))K'_{\jmath_n(y)}(y)
\end{aligned} \tag{A.14}$$

by induction hypothesis. So the two coincide for almost all $y > 0$.

*Induction step for $K_n$.* A straightforward computation shows that the mapping $y \mapsto \frac{c^n(x,y)}{y-x}$ is non-increasing and hence for $\delta > 0$

$$K_n(y+\delta) = \inf_{\zeta\leqslant y+\delta}\frac{c^n(\zeta,y+\delta)}{y+\delta-\zeta} \leqslant \inf_{\zeta\leqslant y}\frac{c^n(\zeta,y+\delta)}{y+\delta-\zeta} \leqslant \inf_{\zeta\leqslant y}\frac{c^n(\zeta,y)}{y-\zeta} = K_n(y)$$

proving that $K_n$ is non-increasing.

Using that $c^n(x,\cdot)$ is non-decreasing and that $\xi_n$ is continuous, local Lipschitz continuity of $K_n$ now follows from

$$K_n(y) \leqslant \frac{c^n(\xi_n(y+\delta),y)}{y-\xi_n(y+\delta)} \leqslant \frac{c^n(\xi_n(y+\delta),y+\delta)}{y-\xi_n(y+\delta)} = K_n(y+\delta)\left(1 + \frac{\delta}{y-\xi_n(y+\delta)}\right)$$

if $\xi_n(y) < y$ and if $\xi_n(y) = y$, recalling (2.15), we have

$$\begin{aligned}
K_n(y+) = \;& \inf_{\zeta\leqslant(y+)}\left\{\frac{c_n(\zeta) - c_{\imath_n(\zeta;y+)}(\zeta)}{(y+)-\zeta} + K_{\imath_n(\zeta;y+)}(y)\right\} \\
\overset{\text{Lemma 2.13}}{=} \;& \inf_{y\leqslant\zeta\leqslant(y+)}\left\{\frac{c_n(\zeta) - c_{\imath_n(\zeta;y+)}(\zeta)}{(y+)-\zeta} + K_{\imath_n(\zeta;y+)}(y+)\right\} = K_{\imath_n(y;y)}(y+) \\
= \;& K_{\jmath_n(y)}(y+),
\end{aligned}$$

and local Lipschitz continuity of $K_n$ follows by induction hypothesis. Equation (A.3) is also proven.

Local Lipschitz continuity of $c^n(\cdot, y)$ follows from the properties of $\imath_n$, cf. (2.8), the fact that the functions $c_i, i = 1, \ldots, n$, are locally Lipschitz and a similar expansion of terms in the case when $\xi_n(y) = \xi_i(y)$ for some $i < n$ Let $k = \imath_n(\xi_n(y)+; y)$. This also shows that for any $y \geqslant 0$ kinks of $c^n(\cdot, y)$ come from kinks in $c_i, i = 1, \ldots, n$.

In order to prove (A.6) we first exclude all $y \geqslant 0$ such that $\xi_n(y)$ is an atom of $c_n$ and $c_k$ and $\xi_n'(y) > 0$. Amongst all $y \in \{\xi_n' > 0\}$ this is a null-set. By assumption $c^n(\cdot, y)$ is differentiable at $\xi_n(y)$. Recalling the equations (2.34)–(2.36), a direct computation proves (A.6) for $k = \imath_n(\xi_n(y)+; y)$. Now we want to apply the induction hypothesis to $c^k$. By choice of $y$ we have that $c_k$ is differentiable at $\xi_n(y) = \xi_k(y)$, i.e. $\mu_k$ does not have an atom at $\xi_n(y)$. Hence, by the assumption that the embedding for the first $n-1$ marginals is valid we cannot have $\xi_k'(y) = 0$ (except on a null-set because otherwise the embedding would fail). By (A.7), $c^k(\cdot, y)$ therefore has to be differentiable at $\xi_n(y) = \xi_k(y)$. This shows that we can indeed inductively apply (A.6) to deduce for $j = \jmath_n(y)$ and $j$ such that $n > j > \jmath_n(y)$ and $\xi_n(y) = \xi_j(y)$ the following equation,

$$\begin{aligned}
0 &= K_n(y) + c_n'(\xi_n(y)) - c_k'(\xi_n(y)) + K_k(y) \\
&= K_n(y) + c_n'(\xi_n(y)) - c_k'(\xi_n(y)) + c_k'(\xi_n(y)) - c_j'(\xi_n(y)) + K_j(y) \\
&= K_n(y) + c_n'(\xi_n(y)) - c_j'(\xi_n(y)) + K_j(y).
\end{aligned}$$

Equation (A.6) is proven.

For later use we note the equation

$$\frac{c_n(\xi_n(y)) - c_{\jmath_n(y)}(\xi_n(y))}{y - \xi_n(y)} + K_{\jmath_n(y)}(y) = K_n(y) = \frac{c_n(\xi_n(y)) - c_k(\xi_n(y))}{y - \xi_n(y)} + K_k(y) \tag{A.15}$$

for $k$ such that $n > k > \jmath_n(y)$ and $\xi_n(y) = \xi_k(y)$.

Finally, we prove the claimed ODE for $K_n$ in the case $\xi_n(y) < y$. For almost all $y \geqslant 0$ we have

$$\begin{aligned}
K_n'(y) &= \lim_{\delta \to 0} \frac{1}{\delta} \left[ \frac{c^n(\xi_n(y+\delta), y+\delta)}{y + \delta - \xi_n(y+\delta)} - \frac{c^n(\xi_n(y), y)}{y - \xi_n(y)} \right] \\
&= \lim_{\delta \to 0} \frac{1}{\delta} \left[ \left( \frac{1}{y + \delta - \xi_n(y+\delta)} - \frac{1}{y - \xi_n(y)} \right) c^n(\xi_n(y+\delta), y+\delta) \right. \\
&\qquad \left. + \frac{c^n(\xi_n(y+\delta), y+\delta) - c^n(\xi_n(y), y)}{y - \xi_n(y)} \right] \\
&= \frac{\xi_n'(y) - 1}{y - \xi_n(y)} K_n(y) + \frac{1}{y - \xi_n(y)} \left( \lim_{\delta \to 0} \frac{c^n(\xi_n(y+\delta), y+\delta) - c^n(\xi_n(y), y)}{\delta} \right).
\end{aligned}$$

The main technical difficulty comes from the possibility that $\xi_n(y) = \xi_k(y)$ for some $k < n$. We present the arguments for this case and leave the other (much easier) case, to the reader.

By assumption the last limit exists and hence we can compute it using some "convenient" sequence $\delta_m \downarrow 0$ where $\delta_m$ is such that $\jmath_n(y+\delta_m) = l$ for all $m \in \mathbb{N}$. Recall (2.9). For $\delta_m$ small enough such that $\imath_n(\xi_n(y); y+\delta_m) = \jmath_n(y)$ we obtain

$$\begin{aligned}
& c^n(\xi_n(y+\delta_m), y+\delta_m) - c^n(\xi_n(y), y+\delta_m) \\
= \; & c_n(\xi_n(y+\delta_m)) - c_l(\xi_n(y+\delta_m)) + (y+\delta_m - \xi_n(y+\delta_m))K_l(y+\delta_m) \\
& -c_n(\xi_n(y)) + c_{\jmath_n(y)}(\xi_n(y)) - (y+\delta_m - \xi_n(y))K_{\jmath_n(y)}(y+\delta_m) \\
\overset{(A.15)}{=} & c_n(\xi_n(y+\delta_m)) - c_l(\xi_n(y+\delta_m)) + (y+\delta_m - \xi_n(y+\delta_m))K_l(y+\delta_m) \\
& -c_n(\xi_n(y)) + c_l(\xi_n(y)) - (y - \xi_n(y))(K_l(y) - K_{\jmath_n(y)}(y)) \\
& -(y+\delta_m - \xi_n(y))K_{\jmath_n(y)}(y+\delta_m).
\end{aligned}$$

From this we obtain for almost all $y \geqslant 0$ by using the induction hypothesis

$$\begin{aligned}
& \lim_{m\to\infty} \frac{c^n(\xi_n(y+\delta_m), y+\delta_m) - c^n(\xi_n(y), y+\delta_m)}{\delta_m} \\
= \; & \xi_n'(y+)\Big[c_n'(\xi_n(y)+) - c_l'(\xi_n(y)+) - K_l(y)\Big] \\
& +K_l(y) + (y-\xi_n(y))K_l'(y) - K_{\jmath_n(y)}(y) - (y-\xi_n(y))K'_{\jmath_n(y)}(y) \\
\underset{(A.2)}{\overset{(A.6)}{=}} & -\xi_n'(y+)K_n(y).
\end{aligned} \tag{A.16}$$

Together with (A.5) this yields in the case when $c^n(\cdot, y)$ is differentiable at $\xi_n(y)$

$$\begin{aligned}
K_n'(y) &= \frac{\xi_n'(y) - 1}{y - \xi_n(y)} K_n(y) + \frac{1}{y-\xi_n(y)}\left(-K_n(y)\xi_n'(y) + \frac{\partial c^n}{\partial y}(\xi_n(y), y)\right) \\
&= -\frac{K_n(y)}{y-\xi_n(y)} + \frac{1}{y-\xi_n(y)}\Big(K_{\jmath_n(y)}(y) + (y-\xi_n(y))K'_{\jmath_n(y)}(y)\Big).
\end{aligned} \tag{A.17}$$

In order to finish the proof we just have to establish that (A.17) also holds in the case when $c^n(\cdot, y)$ is not differentiable at $\xi_n(y)$.

To this end, we first argue that (A.16), and hence (A.17), remains true in the case when $c^n(\cdot, y)$ is not differentiable at $\xi_n(y)$, but when the slope of the supporting tangent to $c^n(\cdot, y)$ at $\xi_n(y)$ which passes the $x$-axis at $y$ equals the right-derivative of $c^n(\cdot, y)$ at $\xi_n(y)$. In that case, denoting $k = \imath_n(\xi_n(y)+; y)$,

$$c_n'(\xi_n(y)+) - c_k'(\xi_n(y)+) - K_k'(y) = -K_n(y). \tag{A.18}$$

Recall the sequence $(\delta_m)$. We do not necessarily have $k = \jmath_n(y+\delta_m) = l$. Nevertheless, we argue that

$$c_n'(\xi_n(y)+) - c_l'(\xi_n(y)+) - K_l(y) = -K_n(y) \tag{A.19}$$

holds. We can safely assume that $\xi_n'(y+) > 0$ (in the other case the conclusion of (A.17) remains true). Also it is enough to consider the case when $k > \jmath_n(y+\delta)$ for

all $\delta > 0$ sufficiently small (otherwise we may consider an alternative sequence $(\delta_m)$ where $k = \jmath_n(y + \delta_m)$ for all $m$ and (A.19) would follow from (A.18)). Consequently, $\xi_k'(y+) \geqslant \xi_n'(y+) > 0$. Then, since by induction hypothesis (A.2) holds true for $k$, we must have, for almost all $y$ that

$$c_k'(\xi_n(y)+) - c_j'(\xi_n(y)+) - K_j(y) = -K_k(y) \tag{A.20}$$

holds for some $j \geqslant l$. We can iterate this argument until $j = l$. Then, combining these results we conclude by (A.18) and (A.20) that indeed (A.19) holds.

Now we consider the case of non-smoothness of $c^n(\cdot, y)$ at $\xi_n(y)$ and where $y$ is such that the slope of the supporting tangent to $c^n(\cdot, y)$ at $\xi_n(y)$ which crosses the $x$-axis at $y$ does not equal the right-derivative of $c^n(\cdot, y)$ at $\xi_n(y)$. We show that in this case we have for sufficiently small $\delta > 0$,

$$\xi_n(y) = \xi_n(y + \delta) \qquad \text{and hence} \qquad \xi_n'(y+) = 0, \tag{A.21}$$

which implies that (A.17) holds as well.

To achieve this we place a suitable tangent to $c^n(\cdot, y)$ at $\xi_n(y)$. Since, by assumption, $c^n(\cdot, y)$ has a kink at $\xi_n(y)$ we have some flexibility to do that. Recalling the tangent interpretation of (2.14) we know by choice of $\xi_n(y)$ that we can place a supporting tangent to $c^n(\cdot, y)$ at $\xi_n(y)$ which passes through the $x$-axis at $y$. Alternatively, by choice of $y$, we can place a tangent to $c^n(\cdot, y)$ at $\xi_n(y)$ which crosses the $x$-axis at some $y + \delta > y$. This implies that

$$\arg\min_{\zeta \leqslant y} \frac{c^n(\zeta, y)}{y + \delta - \zeta} = \xi_n(y). \tag{A.22}$$

Assume first that $\xi_n(y) < y$. Denote $k = \imath_n(\xi_n(y)+; y)$. For simplicity of the argument let us also assume that $\xi_n(y) = \xi_k(y) \neq \xi_i(y)$ for all $i \neq k, n$. Also denote $j = \jmath_n(y) = \imath_n(\xi_n(y); y)$.

Now we will use (A.22) to deduce (A.21). By continuity and monotonicity of $\xi_n$ we have for $\delta > 0$ small enough that $\xi_n(y) \leqslant \xi_n(y + \delta) < \xi_n(y) + \epsilon < y$ for some $\epsilon = \epsilon(\delta) > 0$. By taking $\delta$ small enough we can also assume that $k = \max_{\zeta \leqslant \xi_n(y)+\epsilon} \imath_n(\zeta; y + \delta)$. Then we have

$$\inf_{\zeta \leqslant y+\delta} \frac{c^n(\zeta, y + \delta)}{y + \delta - \zeta} \geqslant \inf_{\xi_n(y) \leqslant \zeta < \xi_n(y)+\epsilon} \frac{c^n(\zeta, y)}{y + \delta - \zeta} + \inf_{\xi_n(y) \leqslant \zeta \leqslant \xi_n(y)+\epsilon} \frac{c^n(\zeta, y + \delta) - c^n(\zeta, y)}{y + \delta - \zeta} \tag{A.23}$$

As for the first infimum in (A.23) we know from (A.22) that it is attained at $\zeta = \xi_n(y)$. Now we will show that the second infimum in (A.23) is also attained at $\zeta = \xi_n(y)$. To this end consider the following estimate for $\zeta > \xi_n(y)$,

$$\begin{aligned}
&c^n(\zeta, y + \delta) - c^n(\zeta, y) \\
=&-c_{\imath_n(\zeta; y+\delta)}(\zeta) + c_{\imath_n(\zeta; y)}(\zeta) - (y - \zeta) K_{\imath_n(\zeta; y)}(y) + (y + \delta - \zeta) K_{\imath_n(\zeta; y+\delta)}(y + \delta) \\
\geqslant&-(y - \zeta) K_{\imath_{\imath_n(\zeta; y)}(\zeta; y)}(y) + (y + \delta - \zeta) K_{\imath_n(\zeta; y+\delta)}(y + \delta) \\
\geqslant&-(y - \zeta) K_j(y) + (y + \delta - \zeta) K_j(y + \delta) =: l(\zeta, y; \delta).
\end{aligned} \tag{A.24}$$

Since $l(\cdot, y; \delta)$ is non-decreasing and non-negative, we deduce that

$$\operatorname*{arg\,min}_{\xi_n(y) \leqslant \zeta < \xi_n(y)+\epsilon} \frac{l(\zeta, y; \delta)}{y+\delta-\zeta} = \xi_n(y).$$

Finally, because at $\zeta = \xi_n(y)$ there is equality in (A.24) we can conclude

$$\operatorname*{arg\,min}_{\zeta \leqslant y+\delta} \frac{c^n(\zeta, y+\delta)}{y+\delta-\zeta} = \xi_n(y)$$

as required.

In the case when $\xi_n(y) = y$ we obtain by (2.9) for $\delta > 0$ sufficiently small that $\imath_n(y; y+\delta) = \imath_n(y; y)$ and hence

$$\begin{aligned} \operatorname*{arg\,min}_{\zeta < y+\delta} \frac{c^n(\zeta, y+\delta)}{y+\delta-\zeta} &= \operatorname*{arg\,min}_{y \leqslant \zeta < y+\delta} \left( \underbrace{\frac{c_n(\zeta) - c_{\imath_n(\zeta; y+\delta)}(\zeta)}{y+\delta-\zeta}}_{\geqslant 0} + \underbrace{K_{\imath_n(\zeta; y+\delta)}(y+\delta)}_{\geqslant K_{\imath_n(y;y+\delta)}(y+\delta)} \right) \\ &\overset{(2.15)}{=} \xi_n(y) = y. \end{aligned}$$

The proof is complete. $\square$

# References


[Azéma et al., 1980] Azéma, J., Gundy, R., and Yor, M. (1980). Sur l'intégrabilité uniforme des martingales continues. *Séminaire de probabilités de Strasbourg*, 14:53–61.

[Azéma and Yor, 1979] Azéma, J. and Yor, M. (1979). Une solution simple au problème de skorokhod. *Séminaire de Probabilités XIII: 06, 90-115, LNM 721.*

[Brown et al., 2001a] Brown, H., Hobson, D., and Rogers, L. (2001a). The maximum maximum of a martingale constrained by an intermediate law. *Probab. Theory Relat. Fields*, 119(4):558–578.

[Brown et al., 2001b] Brown, H., Hobson, D., and Rogers, L. C. G. (2001b). Robust hedging of barrier options. *Mathematical Finance*, 11(3):285–314.

[Carraro et al., 2012] Carraro, L., El Karoui, N., and Obłój, J. (2012). On Azéma-Yor processes, their optimal properties and the Bachelier-Drawdown equation. *Annals of Probability*, 40(1):372–400.

[Cox and Obłój, 2011a] Cox, A. M. G. and Obłój, J. (2011a). Robust hedging of double no-touch barrier options. *Finance and Stochastics*, 15(3):573–605.

[Cox and Obłój, 2011b] Cox, A. M. G. and Obłój, J. (2011b). Robust hedging of double touch barrier options. *SIAM Journal on Financial Mathematics*, 2:141–182.

[Cox and Wang, 2013] Cox, A. M. G. and Wang, J. (2013). Root's barrier: Construction, optimality and applications to variance options. *Annals of Applied Probability*, 23(3):859–894.

[Galichon et al., 2013] Galichon, A., Henry-Labordère, P., and Touzi, N. (2013). A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Annals of Applied Probability, to appear*.

[Henry-Labordère et al., 2013] Henry-Labordère, P., Obłój, J., Spoida, P., and Touzi, N. (2013). Maximum maximum of martingales given marginals. *preprint*.

[Hobson, 2010] Hobson, D. (2010). The skorokhod embedding problem and model-independent bounds for option prices. *Paris-Princeton Lectures on Mathematical Finance, Springer, LNM 2003*.

[Hobson and Neuberger, 2012] Hobson, D. and Neuberger, A. (2012). Robust bounds for forward start options. *Math. Finance*, 22(1):31–56.

[Hobson, 1998] Hobson, D. G. (1998). Robust hedging of the lookback option. *Finance Stoch.*, 2(4):329–347.

[Madan and Yor, 2002] Madan, D. B. and Yor, M. (2002). Making markov martingales meet marginals: With explicit constructions. *Bernoulli*, 8(4):509–536.

[Monroe, 1972] Monroe, I. (1972). On embedding right continuous martingales in Brownian motion. *Ann. Math. Statist.*, 43:1293–1311.

[Obłój, 2004] Obłój, J. (2004). The Skorokhod embedding problem and its offspring. *Probab. Surv.*, 1:321–390 (electronic).

[Obłój, 2010] Obłój, J. (2010). Skorokhod Embedding. *Encyclopedia of Quantitative Finance (R. Cont, ed.)*, pages 1653–1657.

[Rogers, 1989] Rogers, L. C. G. (1989). A guided tour through excursions. *Bull. London Math. Soc.*, 21:305–341.

[Skorokhod, 1965] Skorokhod, A. V. (1965). *Studies in the theory of random processes*. Translated from the Russian by Scripta Technica, Inc. Addison-Wesley Publishing Co., Inc., Reading, Mass.

[Strassen, 1965] Strassen, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36(2):423–439.